

Semantic Web Technologies in the Provision of Personalised Patient Education

Susan Quinn (MSc)

Faculty of Computing, Engineering and the Built Environment of the
University of Ulster

Thesis submitted for the degree of Doctor of Philosophy
February 2018

I confirm that the word count of this thesis is less than 100,000 words.

Table of Contents

List of Figures	vii
List of Tables	x
Acknowledgements	xi
Abstract	xii
Abbreviations	xiii
Note on Access to Contents	xv
Chapter 1	1
1.1 Engaging Patients with Health Management	1
1.1.1 Patient Skills to Use Health Information	2
1.1.2 The Use of Information Technology in Healthcare	4
1.2 Research Issues	6
1.3 Research Aim and Objectives	7
1.4 Contributions to Knowledge	9
1.5 Thesis Structure	9
1.6 Research Outputs	12
Chapter 2	14
2.1 Introduction	14
2.2 Health Information Seeking	14
2.2.1 Patient Education	14
2.2.2 Online Health Information Seeking	16
2.2.2.1 Associated Issues	16
2.2.2.2 Online Health Information Seeking Behaviours	17
2.2.3 Health Literacy	19
2.2.3.1 Obtaining and Processing Health Information	20
2.2.3.2 Seeking Health Information Online	21
2.2.3.3 Ehealth Interventions	22

2.2.4	Summary	22
2.3	Personalisation.....	23
2.3.1	Personalised Healthcare Applications	24
2.3.2	Implicit and Explicit Personalisation.....	26
2.3.3	User Modelling	30
2.3.4	Context.....	32
2.3.5	Summary	34
2.4	Ontologies	35
2.4.1	Ontology Representation Languages & Reasoning Technologies	36
2.4.2	Ontology Engineering.....	41
2.4.3	Biomedical and Health Ontologies	42
2.4.4	Knowledge-Driven Approaches to Ehealth Literacy Applications	44
2.4.5	Summary	45
2.5	Conclusion.....	46
Chapter 3	48
3.1	Introduction	48
3.2	Related Work.....	49
3.2.1	Online Health Information Seeking and Uncertainty Management	49
3.2.2	Health Literacy and Ehealth Literacy	51
3.2.2.1	Online Health Information Seeking Behaviour	52
3.3	Methodology	54
3.3.1	Participants.....	55
3.3.2	Health Literacy	55
3.3.3	Ehealth Literacy	56
3.3.4	Health Questions.....	56
3.3.5	Statistical Analysis.....	61
3.4	Results	61

3.4.1	Health Literacy, Ehealth Literacy and Health Questions	61
3.4.2	Health Information Seeking Behaviour	67
3.4.2.1	Time Spent on SERPs and Information Pages.....	68
3.4.2.2	Health Question Answers	74
3.4.2.3	Search Queries	75
3.4.2.4	Accessing Search Results	77
3.5	Discussion	78
3.5.1	Study Limitations.....	83
3.6	Conclusion.....	83
3.6.1	Recommendations for Future Research.....	85
3.7	Future Work	86
Chapter 4	88
4.1	Introduction	88
4.2	Related Work.....	90
4.2.1	Web-Based Applications for Diabetic Patients	90
4.2.2	Ontological Models of Diabetes	92
4.3	A Web-Based Architecture for Personalised Patient Education	94
4.4	An Ontological Model for Personalised Diabetic Patient Education	96
4.4.1	A Methodology for Ontology Design.....	96
4.4.2	User Modelling	99
4.4.3	Associations Between the Ontological Models	104
4.5	Rule-Based Personalisation.....	108
4.6	Technical Components of the Web-Based Architecture	114
4.7	Testing	114
4.8	Discussion	121
4.8.1	Issues Associated with Engineering the Patient Education Ontology	122
4.9	Conclusion and Future Work	122

Chapter 5	124
5.1 Introduction	124
5.2 Related Work.....	125
5.3 Methodology for Domain Expert Evaluation	127
5.3.1 Design of an Ontology Modelling Screencast	128
5.3.2 Design of the Ontology Engineering Tasks	130
5.3.3 Profile of Invited Contributors.....	133
5.4 Results of Domain Expert Evaluation	133
5.4.1 Participants.....	133
5.4.2 Analysis of the Knowledge Contributions from Ontology Engineering Tasks	137
5.4.2.1 Symptoms, Treatments and Complications	138
5.4.2.2 Physical Activities	140
5.4.2.3 Patient Health Characteristics	141
5.4.2.4 Summary.....	142
5.4.3 Participant Engagement with the Ontology Engineering Tasks	142
5.4.3.1 Completion of the Technical Tasks	143
5.5 Application-Based Evaluation.....	147
5.5.1 Test Profile.....	148
5.5.2 Results of Application-Based Evaluation	148
5.6 Discussion	150
5.6.1 Study Limitations.....	152
5.7 Conclusion.....	153
5.7.1 Issues Associated with Engineering Health Ontologies	154
5.8 Future Work	155
Chapter 6	157
6.1 Introduction	157

6.2	Related Work.....	159
6.3	Methodology	162
6.3.1	Participants.....	162
6.3.2	Diabetic Patient Education.....	164
6.3.2.1	Generic Patient Education Booklet.....	165
6.3.2.2	Web-Based Personalised Patient Education	166
6.3.3	Eye Tracking.....	169
6.3.4	User Engagement Scale (UES).....	169
6.3.5	Health Questions.....	170
6.3.6	Statistical Analysis.....	171
6.4	Results	171
6.4.1	Engagement with the Generic Booklet and the Personalised Education .	172
6.4.1.1	User Engagement Scale	172
6.4.1.2	Gaze Behaviour.....	174
6.4.1.3	Static Images.....	175
6.4.1.4	Text Content	178
6.4.2	Gaze Pattern Analysis for the Experimental Group.....	179
6.4.2.1	Personalised and Static Images.....	179
6.4.2.2	Personalised Components of the Web-Based Education.....	182
6.5	Discussion	184
6.5.1	Engagement Behaviour of the Control Group and Experimental Group Members.....	184
6.5.2	Attention Towards the Web-Based Personalised Education	187
6.5.3	Study Limitations.....	189
6.6	Conclusion.....	189
Chapter 7	192
7.1	Introduction	192

7.2	Research Objectives	192
7.3	Summary of Knowledge Contributions.....	196
7.4	Limitations.....	201
7.5	Future Developments	202
7.6	Recommendations for Similar Research Studies	205
7.7	Recommendations for Health Information Provision.....	206
7.8	Conclusion.....	208
	References	211
	Appendix A	258
	Appendix B	260
	Appendix C	262
	Appendix D	263
	Appendix E	267

List of Figures

Fig. 2.1. Taxonomical model of ontology sub-tree which represents diabetes medications.	36
Fig. 3.1. Protocol for online health information seeking experiment.	55
Fig. 3.2. HCI Browser and HCI Parser components, (a) screen capture of HCI Browser interface for submitting a health question answer, (b) log file containing records for web page visit, health question answer and post-task question choice, and (c) excerpt of spreadsheet record for health question including duration, answer, search queries and first web page visited during searching.	60
Fig. 3.3. Bar charts indicating frequencies of (a) NVS scores and, (b) eHEALS scores.	62
Fig. 3.4. Scatterplot with linear regression line illustrating correlation between NVS scores and eHEALS scores.	63
Fig. 3.5. Frequencies of health question scores for study population.	65
Fig. 3.6. Scatterplots with linear regression line illustrating correlations between (a) question score and NVS score, (b) question score and qualification level, and (c) question score and eHEALS score.	66
Fig. 3.7. Total submitted and correct answers for health questions.	67
Fig. 3.8. Boxplots of completion times for correct and incorrect answers, including p-values from t-test comparison.	68
Fig. 3.9. Pie charts indicating (a) percentage of total time spent on visits to different categories of web resource, and (b) percentage of total distinct visits to different categories of web resource.	73
Fig. 4.1. Architecture to provide web-based personalised patient education. (1) Spring MVC handles communication with the client, (2) processing logic handles interactions between server-side components, (3) domain knowledge is represented in an ontology, (4) a reasoning engine is utilised to execute personalisation rules, (5) suitable content is retrieved from the educational repository, and (6) delivered to the patient as a web page.	95
Fig. 4.2. Stages of engineering for the patient education ontology. Design and implementation phases are grey and evaluation stages are yellow.	97
Fig. 4.3. An excerpt of the Patient Model illustrating (a) classes, (b) datatype properties, and (c) object properties.	101

Fig. 4.4. Overview of the Patient Model, indicating main profile classes (grey) and associated asserted datatype (forked arrows) and object (boxes) properties.	103
Fig. 4.5. Association characteristics of the Patient Model, Educational Content Model and Activity Model.	104
Fig. 4.6. Subsumption relationships within the Medical Conditions Model (a) treatments, (b) symptoms, and (c) complications.	107
Fig. 4.7. Structure of a SWRL rule [428].	108
Fig. 4.8. Patient Model instance with object and datatype properties (asserted and inferred).	116
Fig. 4.9. Screen shot of personalised diabetic education for older female patient with eyesight impairment.	117
Fig. 4.10. Screen shot of personalised and contextualised diabetic education for ... low active male patient.	119
Fig. 4.11. Screen shot of personalised education for female diabetic patient with advanced health literacy and eHealth literacy competencies.	120
Fig. 5.1. Excerpts from online multimedia presentation focusing on (a) superclass-subclass relationships, (b) subsumption relationships with multiple layers, and (c) relationships between unrelated ontology classes.	130
Fig. 5.2. Screen shots of ontology engineering tasks that focus on (a) subsumption relationships, (b) classifying terms within an ontological model, (c) assigning cardinality, and (d) classification of classes within an ontological model.	132
Fig. 5.3. Distribution of information technology expertise amongst study participants.	134
Fig. 5.4. Distribution of expertise for diabetes (blue) and obesity (grey) amongst study participants.	135
Fig. 5.5. Word cloud visualisation of terms contributed through the ontology engineering tasks.	136
Fig. 5.6. Pie chart indicating percentage of total knowledge contributions for survey sections.	137
Fig. 5.7. Subsumption relationships within the enhanced physical activity model, including classes to represent (a) light intensity activities, (b) moderate intensity activities, and (c) vigorous intensity activities.	141
Fig. 5.8. Percentage of participants that completed tasks, organised by sub-groups defined by previous technical experience of using an ontology (dark grey), a	

programming or scripting language (light grey), a database (blue), and a mark-up language (black).	146
Fig. 5.9. Averaged contributions related to expertise rating categories for diabetes (light blue), obesity (grey), and information technology (black).	147
Fig. 5.10. Screen shot of personalised diabetic education for male patient which includes information focused on asserted and inferred treatments.	149
Fig. 6.1. Protocol for eye tracking experiment.	162
Fig. 6.2. Information technology expertise ratings for control and experimental groups.	164
Fig. 6.3. Diabetes expertise ratings for control and experimental groups.	164
Fig. 6.4. Portion of a page from the generic education booklet [464] which focuses on physical activity.	166
Fig. 6.5. Structure of web-based personalised education indicating subject matter of the five web pages.	166
Fig. 6.6. Screen shot of education web page that focuses on physical activity, and contains static, personalised and contextualised content.	168
Fig. 6.7. Arrangement of Tobii X60 Eye Tracker and screen during an experiment.	169
Fig. 6.8. Boxplots of the scores for aesthetics (AE scale) and perceived usability (PUs scale) for the generic booklet and web-based personalised education.	173
Fig. 6.9. Bar chart indicating percentage of total fixations on text, static images and personalised images for the personalised education, and percentage of total fixations on text and static images for the generic booklet.	175
Fig. 6.10. Boxplots indicating total number of fixations on static images for experimental and control group.	176
Fig. 6.11. Boxplots indicating total fixation times (in seconds) on static images for experimental and control group.	177
Fig. 6.12. Visualisation of fixations on personalisation education for a member of the experimental group, (a) gaze plot indicating gaze path, and (b) heat map indicating quantity of fixations on different education components.	179
Fig. 6.13. Boxplots indicating total fixations on personalised and static images for experimental group.	180
Fig. 6.14. Boxplots indicating total fixation times (in seconds) on personalised and static images for experimental group.	181

List of Tables

Table 3.1. Health questions presented to a participant during an online health information seeking experiment.....	57
Table 3.2. Results of multiple regression for NVS scores.	63
Table 3.3. Results of multiple regression for health question scores.	64
Table 3.4. Results of t-test comparison of the completion times for the health questions, indicating t statistic and p-value.	69
Table 3.5. Results of t-test comparison of SERPs visit duration during health questions, indicating t statistic and p-value.	69
Table 3.6. Results of t-test comparison of information page visit duration during health questions, indicating t statistic and p-value.....	71
Table 3.7. Ten most frequently visited domains throughout the study, with brackets indicating rank order of total visits and use as an answer source.	74
Table 3.8. Twenty most frequently occurring search query terms in rank order, including synonyms and misspellings of term.	77
Table 4.1. An excerpt of the key profile classes in the Patient Model.	102
Table 4.2. Subset of SWRL rules used in Personalisation and Contextualisation layer.	112
Table 5.1. Accumulated knowledge contributions by subject, with brackets indicating percentage of total contributions.....	139
Table 5.2. Ten most frequently suggested patient health characteristics and associated cardinalities, with brackets indicating rank order.	142
Table 5.3. Completion rate of ontology engineering tasks, with brackets indicating percentage of participants from discipline group.....	144
Table 6.1. Health questions presented to a participant during an eye tracking experiment.....	171
Table 6.2. Fixation metrics for the personalised images included in the web-based personalised education.	183

Acknowledgements

I would like to express my sincere thanks and appreciation to my Supervisor team of Dr Raymond Bond and Prof. Chris Nugent. Their guidance, support and encouragement has been an invaluable aspect of my PhD experience. It has been a pleasure to work with you both.

I would also like to thank my family, my mother, brothers Jim and Francis, sister Mairead, and Conor. Their unwavering belief in me has been a continuous support at every step of my PhD journey.

Abstract

Health information is a valuable resource for citizens which can assist them to navigate healthcare services, and make medical and lifestyle decisions. Increased Internet availability has facilitated convenient access to a wide range of health information resources. However, whilst health knowledge has become increasingly available, many individuals face obstacles when attempting to engage with the information obtained. Research challenges exist in finding strategies to formulate health education that will enhance its usability and efficacy for all citizens. This thesis considers the human factors associated with utilising health information, and the technological challenges associated with delivering health information. This investigation was fulfilled through the completion of four research studies.

The research initially concentrated on the use of online health information. The first study investigated the online health information seeking behaviours of health consumers, and determined whether these behaviours could be associated with an individual's health literacy and eHealth literacy skills. The focus then shifted to consider the characteristics of generic patient education and it was proposed that a novel personalisation strategy could enhance the usability and attractiveness of the education. The second study focused on the development of a web-based architecture that created personalised education for diabetic patients. Semantic web technologies, including an ontology and a rule-based personalisation component, were incorporated into the architecture. Subsequently the third study concentrated on the ontological knowledge base and sought to strengthen the validity of this knowledge model. A novel methodology for collaborative evaluation of the ontology was described and evaluated. The final study examined user engagement with different formats of patient education. The study compared the engagement behaviours of individuals that were using either a generic education booklet or electronically generated personalised patient education.

The thesis proposes that appreciating the characteristics and aptitudes of citizens can assist with developing effective digital health information services.

Abbreviations

AE	Aesthetics
AOI	Area of Interest
BMI	Body Mass Index
eHEALS	The eHealth Literacy Scale
EU	European Union
IDF	International Diabetes Federation
JSP	JavaServer Pages
NHS	National Health Service
NVS	Newest Vital Sign
OHISB	Online Health Information Seeking Behaviour
OWL	Web Ontology Language
PUs	Perceived Usability
RDF	Resource Description Framework
RDFS	RDF Schema
SD	Standard Deviation
SERP	Search Engine Results Page
SMOG	Simple Measure Of Gobbledygook
SNOMED CT	Systematized Nomenclature of Medicine-Clinical Terms
SPARQL	SPARQL Protocol and RDF Query Language
SWRL	Semantic Web Rule Language
UES	User Engagement Scale
UK	United Kingdom
UMT	Uncertainty Management Theory
URL	Uniform Resource Locator

US	United States of America
WHO	World Health Organization
XML	Extensible Mark-Up Language

Note on Access to Contents

"I hereby declare that with effect from the date on which the thesis is deposited in Research Student Administration of Ulster University, I permit

1. the Librarian of the University to allow the thesis to be copied in whole or in part without reference to me on the understanding that such authority applies to the provision of single copies made for study purposes or for inclusion within the stock of another library.
2. the thesis to be made available through the Ulster Institutional Repository and/or EThOS under the terms of the Ulster eTheses Deposit Agreement which I have signed.

IT IS A CONDITION OF USE OF THIS THESIS THAT ANYONE WHO CONSULTS IT MUST RECOGNISE THAT THE COPYRIGHT RESTS WITH THE AUTHOR AND THAT NO QUOTATION FROM THE THESIS AND NO INFORMATION DERIVED FROM IT MAY BE PUBLISHED UNLESS THE SOURCE IS PROPERLY ACKNOWLEDGED".

Chapter 1

Introduction

1.1 Engaging Patients with Health Management

The role and responsibilities of a patient within a healthcare system have changed in recent decades. Earlier models of healthcare were characterised by a paternalistic format whereby the physician acted to make decisions about their patients' needs and instructed the patient accordingly [1, 2]. However, recent models of healthcare provision have advocated for greater patient engagement with managing their own healthcare and participation in decisions about their care [3-8]. For example, many of the recent health frameworks in the United Kingdom (UK) emphasise that patients should be supported in becoming more knowledgeable about how to manage their health, and to gain a sense of confidence and control over their health and wellbeing [5, 9, 10]. Furthermore, patients should be involved in shared health decisions, making informed choices about their treatment [5, 9-11]. Making informed decisions necessitates that patients have access to accurate, relevant and comprehensible information about their health and the services available to them [4, 10-14]. Effective use of health information can positively influence quality of life, as patients with greater health knowledge may be less anxious, can have a better understanding of how to avoid health risks, and tend to adapt better to changes in their health status [9, 15-18]. Moreover, informed patients may also use health services less often than those with reduced health knowledge, having fewer unscheduled hospital admissions and tending to choose less invasive treatments [17, 19, 20].

For many patients, medical professionals are still a principle source of trusted healthcare advice [21-25]. Moreover, patient information leaflets still constitute a key method of distributing evidence based health education to patients [1, 24, 26, 27]. However, the channels from which health information is consumed are diversifying, and health consumers are increasingly gathering information from newspapers and magazines, television and radio programmes, and talking with friends, relatives or peers [21-23]. Moreover the Internet is also a prominent resource for health information seekers [22, 23, 28, 29]. However, gaining access to health information does not ensure that an individual will be able to attain the full benefits of this resource. Primarily the acquired information needs to be relevant and applicable to a patient's health needs, however

even if this is so, the usefulness of the information can be diminished if it is not fully comprehensible for the patient. This may be due to the subject matter under review; health information is complex and the terminology used can be unfamiliar to non-medical personnel. However, equally of relevance are the skills of the patients themselves as, for example, the utility of health information can be reduced if individuals do not possess the requisite cognitive skills to interpret and apply the information.

Health education is a fundamental process in strategies to inform and engage patients with effective health management. Associated research in areas including health communication and health informatics have expanded in recent years, however challenges remain in understanding the behaviours employed to interact with novel channels of health information such as Internet based health applications. This project aims to contribute to research efforts associated with the use of electronic formats of health information. The project focuses on the human behaviours associated with using online health information, and on the technologies used to provide online health information applications. The following two sections will provide further background on how human factors can influence the utility and efficacy of health information, and how information technology is being used to support health service provision.

1.1.1 Patient Skills to Use Health Information

The capability of an individual to make health decisions and manage health problems is very much influenced by their skills to obtain suitable health information and effectively apply this to their health situation. The skills to interpret and use health information is regarded as an individual's *health literacy*. Perceptions of the meaning, determinants and outcomes of being health literate are diverse [30], however one widely cited definition of health literacy is provided by the Institute of Medicine, "*the degree to which individuals have the capacity to obtain, process and understand basic health information and services needed to make appropriate health decisions*" [31]. This description highlights a commonality of most definitions whereby health literacy is conceptualised as a set of personal skills which assist an individual to use health information to make daily and longer term health choices, and supports interactions with healthcare services [30, 32, 33].

Having an adequate level of health literacy can empower an individual to have positive health behaviours. However, several national surveys have highlighted a significant prevalence of limited health literacy amongst populations worldwide. The *European Health Literacy Survey (HLS-EU)* assessed health literacy competency within eight European Union (EU) member states and discovered that nearly half of the respondents had skills that were quantifiably insufficient [34]. Analogous findings have emerged from research in the United States of America (US). The *2003 National Assessment of Adult Literacy (NAAL)* [35] revealed that 36% of American adults had basic or below basic health literacy. Disparities in health literacy proficiency are often associated with personal, socio-economic, and health attributes. The characteristics that are most frequently associated with inadequate health literacy skills include being older [34-38], and having reduced income [34, 36, 37], less education [34, 36-38] and reduced health status [34, 39]. Inadequate health literacy has been linked to negative health practices including a lack of use of disease prevention services, increased health risks, reduced medication compliance and poorer health outcomes including higher mortality rates [36, 40-43]. A prevalence of low health literacy amongst population groups can also increase demand for healthcare resources, placing an increasing financial burden on national healthcare budgets [36, 40, 44-46]. In consequence, it is no surprise that health literacy is gaining increasing attention in both clinical care and health policy research and government health policies [33, 36, 47, 48]. In the US, improving health literacy is a key objective of the *Healthy People 2020* agenda [49], while the *National Action Plan to Improve Health Literacy* [50] aims to involve organisations, policy makers and communities in strategies to enhance health information provision. Several EU health strategies have recognised health literacy as a significant determinant of citizen empowerment in health matters [8, 51], and the World Health Organization (WHO) has recommended increased investment in health literacy research [36].

Various projects have identified a relationship between health literacy and amount of health knowledge [52-55]. Moreover, research findings would also suggest that limited health literacy is an impediment to becoming better informed about health and wellbeing. When compared to patients with adequate health literacy skills, those with reduced skills often find it difficult to process, understand and recall medical information and instructions [56-61]. Comprehending health information is a foundational step towards becoming an informed patient, thereby high levels of limited

health literacy amongst patients can present an obstacle to efforts to increase patient engagement with their own health matters. It has been recommended that healthcare professionals should modify and tailor the information they provide in accordance with the health literacy aptitudes of their patients [62-65]. However, difficulties remain in identifying optimal methods of delivering health information that are efficacious for patients with varying levels of health literacy. Information technology may provide a means to develop digital solutions that can facilitate automated adaptable health information provision. This is a key proposition of this research project and will be discussed in greater detail in subsequent chapters of this thesis.

1.1.2 The Use of Information Technology in Healthcare

As indicated in Section 1.1, the sources used by health consumers to obtain health information are expanding. Most notably, increased Internet access has facilitated expanding use of online health information resources. For example, in a survey of US adult Internet users, health information seeking was found to be the third most popular online activity that is measured [28], with 72% having used the world wide web to search for health information [23]. Research amongst the 28 EU member states revealed comparable behaviours. The results of a 2014 Eurobarometer denoted that 59% of the citizens surveyed had searched online for health-related information, with 10% of this group having done so at least once a week [22]. The Internet provides access to a volume of health resources encompassing a wide range of prevalent and less known health concerns. Governments and medical institutions have also recognised the utility of web-based platforms to deliver authoritative health advice. For example, *NHS Choices* [66], the digital information channel of NHS England, provides cost free health information and health management tools. Since its launch in 2007, traffic to the portal has continuously grown and it currently receives 50 million visits each month [67]. In the US the National Library of Medicine maintains *MedlinePlus* [68], a portal which provides links to free health information on over 1000 topics along with access to a medical encyclopaedia and dictionary. The value of trustworthy online health information has also been recognised in the *Healthy People 2020* agenda, which includes a commitment to increase the quantity of quality healthcare related websites [49].

In addition to providing online information channels, the adoption of information technology in healthcare is a common strategic focus in many countries worldwide. *Ehealth* has been described as “*the use of ICT in health products, services and processes combined with organisational change in healthcare systems and new skills, in order to improve health of citizens, efficiency and productivity in healthcare delivery, and the economic and social value of health. eHealth covers the interaction between patients and health-service providers, institution-to-institution transmission of data, or peer-to-peer communication between patients and/or health professionals*” [69]. A survey by the WHO found that 58% of the 125 responding member states had national strategies for eHealth which included objectives for health information provision and telecare [70]. Ehealth services have been proposed as a method of sustaining healthcare delivery in the face of ever increasing patient demand, and as a means to support a patient-centred approach to care [5, 8, 49, 71]. Electronic services can reduce the burden on frontline services provided by GP surgeries or accident and emergency departments [72], and can also contribute to longer term benefits for healthcare budgets [73]. Providing services online may also increase accessibility and convenience for patients. For example, the UK government has expressed a commitment to increase the availability of online services for booking GP appointments and ordering repeat prescriptions [5, 71, 74]. Ehealth solutions have the potential to enhance patient experience and health outcomes, however in order to use these services, many patients will be required to possess a level of digital aptitude. Consequently, there will be variability amongst patients in the ability to consume and obtain the full benefits from eHealth services. The concept of *digital exclusion* describes the situation of many members of society whereby they do not have sufficient Internet access or adequate digital skills and confidence to engage with eHealth services [72]. In many cases, there is an overlap between being at risk of poor health and being digitally excluded [72, 74]. This has led to concerns that digital exclusion may contribute to health inequalities amongst population groups [72, 75]. Proposed solutions to tackle disparities in digital experience include the development of digital health skills amongst targeted population groups [72]. Moreover, there are calls for developing eHealth solutions that can be usable by health consumers with varying level of digital skills [75, 76]. The use of information technology to support health information provision is a key research focus of this project. Furthermore, it is also proposed that understanding the skills of health

consumers to utilise electronic sources of health information may contribute to developing effective electronic solutions.

1.2 Research Issues

Section 1.1 established the context of this research project by indicating how health knowledge can assist a patient to effectively manage their health. Also described were the various challenges that can exist for patients with relation to becoming informed about their healthcare choices, and in gaining access to digital health tools. Based on this discussion a number of research issues have arisen which are described below.

- Many contemporary healthcare settings adhere to a model of care whereby patients are encouraged to engage with healthcare services and participate in shared decisions about their health. However, making informed health choices can only take place when a patient is knowledgeable about their health. Nevertheless, patients will vary in their capability to understand and apply health knowledge. Subsequently, various challenges exist in realising methods to deliver health information that will improve accessibility and enhance the clarity of the information for patients.
- Several large scale surveys have identified significant levels of inadequate health literacy amongst populations. Having limited health literacy can reduce an individual's aptitude to understand, retain and utilise health information. Accordingly, this may impact a patient's ability to engage with health services, and fulfil their health needs. This suggests that it could be beneficial for patients to receive education that has been adapted as suitable to their health literacy skills. Moreover, formulating health education to attract the attention and interest of the reader may also encourage higher levels of engagement. Achieving these aims may require analysis of novel strategies and technical approaches to present education that is utilisable for a range of patient abilities. Developing electronic solutions will require the use of technologies that can express the complexity of medical data related to the pertinent entities in the health domain. Furthermore, the usability of electronic solutions may be enhanced if they are manageable for patients with different levels of digital experience.

- The application of information technology in healthcare provision is regarded as a means to improve access to health services, enhance patient experience and alleviate the pressure on healthcare budgets. However, although services such as online health information portals are highly utilised, there is variability amongst patients in their technical proficiency to use online tools and services. Providing electronic services that are useful and convenient requires an understanding of the digital skills of patients. Online information platforms are a means to provide immediate assistance for non-emergency situations, therefore it is valuable to understand the behaviours and strategies of information seekers when searching for and using online health information. This may highlight the skills and abilities that are associated with successful online interactions, and may also highlight which sources and formats of online information are attractive for health consumers.

1.3 Research Aim and Objectives

The increased availability of Internet enabled devices in many societies worldwide has facilitated a rise in the use of electronic health information applications. Accordingly, the use of digital platforms to support healthcare provision has also been adopted by many governments and health institutions. Consequently, there is a need to appreciate the actions and behaviours of citizens when interacting with online health information application and services. With relation to these issues, the aim of this research study is as follows.

To use software tracking and user log analysis to appraise the health seeking behaviours and skills involved in utilising and engaging with web-based health information applications, and in turn to develop a computer-enhanced personalised patient education system using ontologies and related technologies.

It is intended that the project aim will be achieved by fulfilling the following objectives.

Objective 1.

To develop an understanding of current research activities focused on the communication channels, strategies and technologies used in the provision of electronic health information. This will involve conducting a comprehensive review of associated literature in order to appreciate the methods by which health information is

communicated, and the human activities associated with obtaining and using health information. The review will also concentrate on technologies and software languages that have been employed in the development of electronic health applications. The review will uncover current research in these areas and the challenges being faced. This will enable the identification of areas suitable for further investigation throughout this research study.

Objective 2.

To use software tracking and log analysis to classify and quantify the behaviours of health information seekers when searching for and using online health information, and assess whether these behaviours are associated with health literacy or eHealth literacy skills.

Objective 3.

To design and implement a web-based technical solution that will provide electronic patient education that is tailored as suitable for the personal health needs and skills of patients.

Objective 4.

To develop a novel approach for collaborative evaluation of an ontology in the health domain, and evaluate whether this approach is accessible for ontology engineering novices.

Objective 5.

To appraise the comprehension and engagement behaviours exhibited by individuals when interpreting generic patient education material and computer-aided personalised patient education material, and evaluate whether there are any similarities or distinctions in engagement patterns between both approaches.

Objective 6.

To assess whether the engagement behaviours exhibited by individuals will vary in response to viewing different aspects of personalisation included in a patient education artefact.

1.4 Contributions to Knowledge

Fulfilling the preceding objectives will result in the following contributions to knowledge within the areas of electronic health information provision, personalisation and semantic web technologies.

- 1) The design and implementation of a methodology to track and quantify the online health information seeking behaviours, and measure the health literacy and eHealth literacy skills of health information seekers.
- 2) The implementation of an extensible web-based architecture, comprising an ontological knowledge base and rule-based personalisation, which will deliver education that has been tailored to the preferences, personal health needs, health literacy and eHealth literacy competencies of individual patients.
- 3) The design and implementation of a method for collaborative ontology evaluation that supports domain experts with variable modelling experience to complete ontology engineering tasks and contribute domain knowledge.
- 4) The design and implementation of a methodology to record subjective and objective measures of engagement with patient education artefacts.

1.5 Thesis Structure

This chapter introduces the context, research aim and objectives of this research study. The subsequent chapters, as summarised below, detail the research completed to fulfil the listed objectives.

Chapter 2 - Health Information Applications: Techniques, Technologies and User Behaviours

This chapter introduces the impetus of this PhD by reviewing literature pertaining to three main themes, (1) health information, (2) personalisation and, (3) semantic web technologies. The topic of health information seeking is introduced by describing several channels by which individuals can acquire health information, and the skills involved in locating and using health information. Subsequently the use of personalisation to enhance the efficacy of health applications is examined. The review

finishes by examining the use of ontologies in healthcare applications, and highlights recent research activities that focus on ontology engineering. For each of the three themes, pertinent advances in research and associated issues are highlighted.

Chapter 3 - Quantifying Health Literacy and eHealth Literacy Using Existing Instruments and Browser-based Software for Tracking Online Health Information Seeking Behaviour

This chapter appraises the information seeking strategies employed by individuals when searching for and using online health information. The chapter describes a study in which 60 participants were recruited to complete online search tasks. A novel methodology for scrutinising online health information seeking is presented. Various health information seeking behaviours are quantified and search task accuracy is measured. Subsequent analysis establishes whether any of the noted online health information seeking behaviours are associated with an individual's level of health literacy and eHealth literacy skills.

Chapter 4 - Ontological Modelling and Rule-Based Reasoning for the Provision of Personalised Patient Education

This chapter evaluates the use of semantic web technologies to create personalised health services. The chapter describes the design and implementation of a web-based architecture that provides personalised education for diabetic patients. An underpinning component is an ontological knowledge base which captures clinical aspects of diabetes, and a user model representation of a diabetic patient. The various stages of ontology engineering are detailed. The development of a rule-based personalisation component is also described. The chapter describes how these components function to output various personalised features within the education experience. Testing of the architecture is depicted by the use of vignettes which illustrate how the produced education can fulfil the personal health needs, competencies and preferences of different patients.

Chapter 5 - A Two-Staged Approach to Evaluating an Ontology for Delivering Personalised Education to Diabetic Patients

This chapter examines collaborative approaches to ontology engineering. The chapter describes the two stages of evaluation applied to the personalised patient education

ontology. The main focus is on the first evaluation phase whereby domain experts with variable levels of technical modelling knowledge were invited to participate in the evaluation of the ontology. The design and realisation of a methodology to assist participants with completing ontology engineering tasks is described. Analysis of the participants' modelling decisions is detailed, along with descriptions of how this knowledge was used to evaluate and enhance the ontology. Subsequently, an evaluation of the viability of involving ontology engineering novices in evaluation processes is provided. The second stage of evaluation incorporated the refined ontology within the architecture to assess whether it functioned to support the provision of personalised patient education.

Chapter 6 - An Evaluation of User Engagement with Generic and Personalised Patient Education for Diabetic Patients

This chapter appraises patterns of human engagement with different formats of patient education. The chapter describes a between subjects study which examined engagement with a patient education booklet authored by a registered UK charity, and education produced by the personalised patient education system developed in this PhD. A methodology for quantifying attributes of engagement is presented. Appraisal of the engagement behaviours of the participants when viewing the educational artefacts is detailed. Subsequent analysis focuses on the personalised education and examines levels of engagement with different personalised features.

Chapter 7 - Conclusion and Scope for Future Work

This chapter describes how the aim and objectives have been fulfilled through the various research studies completed as part of this research project. The manner in which the contributions to knowledge have been achieved, and the conclusions drawn from the research are also detailed. The thesis is completed with a discussion of recommendations for future research directions.

1.6 Research Outputs

The research completed during this PhD has been disseminated through the following conference and journal publications.

Peer-Reviewed Journal Articles

S. Quinn, R. Bond and C. Nugent, "Quantifying health literacy and eHealth literacy using existing instruments and browser-based software for tracking online health information seeking behavior," *Comput. Hum. Behav.*, vol. 69, pp. 256-267, 2017.

S. Quinn, R. Bond and C. Nugent, "Ontological modelling and rule-based reasoning for the provision of personalized patient education," *Expert Syst*, vol. 34, pp. e12134, 2, 2017.

S. Quinn, R. Bond and C. Nugent, "A two-staged approach to developing and evaluating an ontology for delivering personalized education to diabetic patients," *Informatics for Health and Social Care*, pp. 1-16, 2017. doi: 10.1080/17538157.2017.1364246.

Peer-Reviewed Conference Proceedings

R. Bond, S. Quinn and C. Nugent, "The Development of an Ontology to Facilitate Personalised Patient Education," in *Medicine 2.0 Congress, Malaga. IJMI*, 2014.

S. Quinn, R. Bond and C.D. Nugent, "An Ontology Based Approach to the Provision of Personalized Patient Education," in *Ambient Assisted Living and Daily Activities. IWAAL 2014. Lecture Notes In Computer Science, Vol. 8868*, L. Pecchia, L. L. Chen, C. Nugent, and J. Bravo, Eds. Cham: Springer, 2014, pp. 67-74.

S. Quinn, R. Bond and C. Nugent, "A Semantic Computer-Enabled Architecture For The Provision Of Personalised Patient Education," in *7th Annual Translational Medicine Conference, Derry/Londonderry. CTRIC*, 2015.

S. Quinn, R. Bond and C. Nugent, "An investigation into the relationship between health literacy, eHealth literacy and online health information seeking behaviour," in *19th Annual Health Promotion Conference. Health Literacy: Research, Policy and Practice*, 2015.

Software

The first research study, which is described in Chapter 3, included the development of a parsing tool to process online health information seeking log files. This tool was developed using Java and is available to download from the github website [77].

Chapter 2

Health Information Applications: Techniques, Technologies and User Behaviours

2.1 Introduction

Chapter 1 established the context of this research project by highlighting the various challenges that health consumers can encounter when attempting to engage with health information and electronic health tools. This chapter further describes the background to the research by presenting a review of related work, and highlighting current issues. This chapter describes a review of health information seeking behaviour research and the challenges of measuring and improving digital health literacy amongst citizens. In turn, given the proposed hypothesis that personalised patient education could increase patient engagement and their health literacy, this chapter presents a review of personalisation techniques with a focus on knowledge engineering using ontologies.

Section 2.2 addresses health information seeking by health consumers. This includes descriptions of the various sources of health information available, and the methods by which health information is accessed and processed. Section 2.3 describes the use of information technology to provide personalised health applications, in particular through the use of semantic web technologies. The benefits of incorporating personalisation and contextualisation in health applications are described, and moreover, the various challenges to achieving personalisation are also highlighted. Section 2.4 concentrates on the use of ontologies in personalised healthcare applications. This includes the semantic web technologies available for ontology development and examples of ontology-based frameworks. Ontology engineering is also covered, including more recent research directions which focus on collaborative engineering projects.

2.2 Health Information Seeking

2.2.1 Patient Education

Printed patient education materials are often used to augment verbal information provided during visits to physicians, or after surgery [26, 27, 78-80]. Thereby many patients will be required to read and use written health information at the point of need.

Nevertheless, the effectiveness of this educational approach may be hampered due to an individual's inability or motivation to engage with the material. A patient may have difficulty understanding the medical terminology and images used, or may feel overwhelmed, anxious or confused by the volume of information introduced [65, 79, 81]. These factors may result in the individual finding it difficult to identify important points or misinterpreting fundamental information [82, 83]. The readability of written education has been widely investigated. In many cases the materials examined were found to be at a higher readability level than the national recommendation [84-86]. In addition to the words and grammar used, the quality and usefulness of patient education may also be reduced by the inclusion of images that do not supplement the written text or do not draw the interest of the reader [85, 86]. Additional content deficiencies include a lack of detail related to pertinent patient information [87], no clear indication of the purpose of the document [85], and the inclusion of information that did not comply with national recommendations [85, 87].

However, despite these limitations, patient feedback regarding printed materials has been mixed. For many patients the documents produced by health organisations are regarded as an accurate and trustworthy form of health information [24, 65]. Moreover, an investigation of patient experiences of foot surgery information leaflets found that most individuals rated the quality and comprehensibility of the materials very favourably [27]. However, in contrast, other studies have found that the terminology used in educational materials can be challenging for patients to understand, and that some patients would prefer simple and familiar language [65, 79, 86]. Further variation is found with relation to the volume of content included. While some patients have suggested that more detailed information should be included [27], others have expressed a desire to be able to choose between different levels of complexity and detail in the information supplied [78].

Despite the continued availability of printed health materials many patients are increasing the communication channels by which they consume health information. For many information seekers Internet resources serve as a complement to health services and information gathered from physicians, printed media, friends or family [88]. The motivations of online health information seekers will be further described in the next section.

2.2.2 Online Health Information Seeking

Health information seekers are increasingly using online health information to answer health-related questions. The Internet provides a convenient, cost effective and private means of gaining access to health knowledge, and the motivations for seeking information are diverse. Information seekers have used the Internet as a diagnostic tool, seeking to understand symptoms, illnesses and injuries, medications and treatments [22, 23, 89, 90]. Many have also sought advice from online peers, and have spent time reading or watching the health experiences of others, and posting questions, comments and stories about their own health situation [22, 23, 91]. Online health information seeking is often carried out in preparation for a visit to a physician, or afterwards to supplement, clarify, or in some cases challenge the advice given [88, 92-95]. For many individuals health information is sought for reassurance, to lessen uncertainty, and to help them to reconcile with a new health situation [88, 92, 96]. Many individuals are satisfied with the information found, rendering it relevant and useful to their needs [22, 94, 97], moreover increased knowledge has helped some patients to gain a sense of competency and control in their health situation [98, 99].

However, whilst many benefits have been attributed to the use of online health information, doubts have also been raised about the quality, accuracy, reliability and veracity of various online health resources [100-102]. Furthermore, there are concerns that not all health information seekers will have sufficient skills to appraise the quality of online health information [103]. Moreover, as many health consumers admit that they do not always discuss the health information they find with their physicians [23, 97] there is an increased potential for patients to misunderstand health information and to place their trust in unreliable information. For example, a survey by [104] found that 11.2% of the respondents had refused or discontinued treatments prescribed by doctors and dentists after considering online health information. For many information seekers, especially those that have higher levels of health anxiety, online searching may result in increased confusion, fear, and worry [92, 105, 106].

2.2.2.1 Associated Issues

Intention to seek and use Internet based health information has been positively associated with a health consumer's perceived usefulness of online health information [90, 107, 108]. However, even for those that are willing to engage with online resources,

challenges can exist. Searching for health information can be a time-consuming activity and health consumers can be overwhelmed and confused by the volume of information available [22, 88, 109, 110]. For many individuals difficulties are faced in understanding the health information found. The readability of a piece of text refers to the reading comprehension level that a person must have to understand it [111]. Multiple studies have identified online health websites, including those of professional medical associations, that contain written content that exceeds recommended readability levels [112-114]. There are concerns that incomprehensible information may discourage a consumer from seeking medical care [114] or encourage them to utilise less reliable sources [112]. Furthermore, there can also be variability in the accuracy and reliability of the health information found. Studies addressing the quality of online health information have found a volume of websites containing information that was outdated, inaccurate and did not comply with official health advice and medical guidelines [100-102, 115]. Willingness to use online health information has been correlated with a consumer's level of trust in the information found [90], however consumers do not always use the recommended standards to judge the trustworthiness of a health website. In a study of web searching habits, Silence et al. [116] found that website aesthetics and design influenced their subjects' initial impressions of a website, as those with poor design were quickly rejected. A professional appearance was associated with trustworthiness by the participants of [117]. Features such as privacy policy, third party endorsement and advertisements have also been found to affect consumer credibility perceptions [118], while the presence of informational features such as testimonials, statistics and author credentials have also been positively correlated with increased website credibility [119]. However, although there is inconsistency in the factors by which consumers judge website credibility, some of the most consistent findings relate to branded health websites. Health consumers are often found to be wary of commercially oriented websites [22, 116, 120] whilst consistently associating reliability and expertise with the online platforms of reputable medical centres and organisations [88, 116, 117, 119, 120].

2.2.2.2 Online Health Information Seeking Behaviours

The volume of publications focused on health information seeking has increased in recent years [121], however despite the growing research, commercial and governmental interest in the topic there is still little consensus on the constituents of

health information seeking behaviours [121, 122]. A significant amount of research has focused on the strategies used to locate information resources. One frequently employed methodology is an observational study whereby areas of investigation are set and the behaviours exhibited by the subjects are recorded through screen capture, logs, interviews and observation. The quantifiers used to evaluate online health information seeking behaviour can include search accuracy (e.g. the quantity of tasks correctly fulfilled), search strategy (e.g. using a search engine or entering a URL), search queries entered, time spent on search results pages or website visits [123-127].

A search for health information often begins through a search engine [116, 117, 123]. This usually requires the information seeker to formulate a search query to express their information needs. However ill-defined search queries may contribute to deficient search results that do not match the consumer's information needs [125, 128-130]. It has been proposed that search query analysis can highlight how consumers conceptualise and express health concepts, and can be used to assist with content provision and the design of websites and information retrieval systems [131, 132]. One frequently used method for appraising search queries is through analysing web browser logs. This renders the benefit of reflecting the natural behaviours of health consumers during searching activities [133]. Much research has focused on the length and composition of health search queries. Various analysis of web browser logs have determined that medical and health search queries tend to be short (e.g. an average of 2.3 terms per query) [117, 128, 130, 132, 134]. However, the length of search queries leading to a health website aimed at the parents of injured children was much longer at 6.27 terms per query [131]. The authors proposed that this increase was due to a significant proportion of question-type queries, which suggests that visitors to the site had specific information queries. Studies of health information seeking on the MedlinePlus website found that information seekers often used natural language to construct search queries [135, 136]. Health consumers often rely on prior knowledge when constructing search queries [125] however they may lack specialised domain knowledge and can understand and express health concepts in a different manner to health professionals [137-139]. This can contribute to difficulties when searching for health information online [132, 137].

Another prominent area of investigation focuses on activities related to search engine result pages. Nevertheless, there is disparity in the findings. While some studies denoted that participants chose results from the first search results page and only occasionally read beyond this page [117, 130, 140], the authors of [124] reported that an average of 5.4 search pages were viewed by their study participants. Further review of the motivations of health consumers explained why this was so; a number of consumers believed that the first page contained the results most relevant to their search [130, 140], however, in contrast, the subjects of [124] indicated that rank order did not necessarily reflect the aptness or credibility of the results. Variation has also been exhibited in the methods used to appraise search results lists. Whilst some participants have displayed a systematic approach to evaluating search results lists, for example by scanning through the results and considering the headers, summaries and URLs, or looking for specific keywords and source identifiers [124, 141-143], in other studies the subjects appeared to select a result at random from the list [126, 130].

2.2.3 Health Literacy

As highlighted in Chapter 1 an individual's capability to locate, understand and apply health information is very much influenced by their health literacy skills. However, health literacy is a complex concept which comprises a diversity of communicative and cognitive skills. Nutbeam [144] conveyed the multifaceted nature of health literacy by proposing three forms of health literacy; *functional*, *interactive* and *critical* health literacy. Functional health literacy encompasses an individual's competencies in reading, writing and numeracy which enable access and understanding of written health information [144]. Interactive health literacy includes the listening and speaking skills required to process different formats of health resources and communicate effectively with healthcare personnel [144]. Critical health literacy comprises cognitive skills that enable a patient to critically appraise, extract and retain health information and to apply it to one's particular health needs [33]. Critical health literacy skills can empower an individual towards autonomous management or increased control of their health status, for example by undertaking disease prevention strategies [30, 144, 145]. An individual's health literacy skills combine to assist them to function effectively within a clinical care environment, and to exercise health judgements in daily lifestyle. However other definitions have further extended health literacy dimensions to include an individual's ability to affect health outcomes not only for themselves but for their

family or community also [33]. For example being health literate enables one to engage with public health issues, or to contribute to solutions which can affect health literacy at a population level [30, 76].

2.2.3.1 Obtaining and Processing Health Information

In many cases health literacy proficiency is correlated with the level of knowledge that an individual has about their own health [40, 53]. Furthermore, it is often found that patients with lower health literacy can have reduced recollection of medical information. For example, in a study in which patients with hypertension were asked to report their prescribed medications, a large proportion of patients with inadequate health literacy were unable to name any of the medications listed in their medical record [59]. Moreover, research has suggested that the health information needs of lower health literacy persons are not always being fulfilled by medical professionals. In two distinct studies which addressed the perceptions of health information provision for patients receiving cancer treatment, it was found that women with more limited health literacy competencies had higher levels of unfulfilled information needs, lower levels of perceived information provision and less satisfaction with the information provided than those with more adequate levels of health literacy [62, 146].

Disparities in health literacy skills between population groups are often associated with variance in health information seeking behaviour. It has been suggested that individuals with advanced health literacy are more proactive health information seekers [147] and may have access to a wide range of health information resources such as books, magazines, newsletters and Internet resources [14, 35, 37]. In contrast, those with reduced health literacy are more inclined to rely on healthcare personnel and friends as a preferred source of health information [14, 54, 148]. Some patients with inadequate functional health literacy have suggested that a lack of reading skills can dissuade them from accessing written materials [14], moreover the complexity of written materials may also be discouraging for some individuals [149]. Health literacy skills are often correlated with an individual's ability to process and reason upon text and numerical based health materials, such as medication information [56, 150].

Research efforts have also examined the manner in which written health information is read and processed. For example, von Wagner et al. [57] used an interactive information

tool to investigate the relationship between health literacy and the cognitive effort required to process complex health information. A comparison of behaviours found that subjects with more limited health literacy skills sought less information than those with advanced health literacy. Moreover, it was also found that the average amount of time spent reading information links was larger for the lower health literacy group, implying that this group had expended great effort to read the information. Meppelink and Bol [58] investigated the associations between health literacy, attention patterns and recall of health information. They found that an increase in the amount of time spent reading the webpage resulted in improved recall only for the subjects that had adequate health literacy skills. The authors suggest that health literacy proficiency may influence the efficiency with which information can be processed.

2.2.3.2 Seeking Health Information Online

Health consumers of varying health literacy aptitudes are using the Internet to obtain health information, however usage is more frequent amongst adults with more advanced skills [14, 35, 37, 151]. For example, Sarkar et al. [152] found that subjects with reduced health literacy skills were less likely to logon to a health patient portal and used the portal functions to a lesser extent than participants with more advanced skills. Much research has concentrated on how trust in online resources is mediated by health literacy skills, however there is little consensus on the findings. An investigation of the use of nutritional health information found that subjects with adequate and more limited levels of nutritional literacy expressed no differences in trust of online health information [153]. An investigation of health information source usage amongst patients with adequate health literacy skills found mixed attitudes [65]. Whilst some respondents indicated that the Internet could provide helpful and reliable health information, others expressed doubts regarding trustworthiness and perceived that online information could cause worry. A mixed-methods study of health literacy and the evaluation of online health information found that although adults with low health literacy were aware that the quality of online health information was variable, they were less aware of established criteria by which to assess the quality of information and were more likely to use non-established criteria [154].

2.2.3.3 Ehealth Interventions

The use of health-based electronic tools targeted for different health literacy competencies is an increasing research focus [75, 155]. Various strategies have been applied to improve the comprehensibility of health information and promote beneficial health behavioural change. As previously noted, consumers with limited health literacy can have difficulty understanding written information, thereby several interventions have adapted the composition and presentation of text content to enhance its readability and understandability [156, 157]. Multiple modalities such as video, animations, and images are often included to assist with engaging the user [155-159]. A novel approach by [160] utilised an animated virtual nurse in an interactive system which provided information about post-discharge self-care for hospital patients. Less health literate patients found the provided information useful, and the system easy to use. The efficacy of electronic interventions is often measured with relation to improved health knowledge or health behaviours, or clinical quantifiers such as weight, blood pressure or psychological symptoms [155]. However, while in many cases electronic interventions are associated with positive health outcomes, in other cases the results are mixed [75, 155, 158].

2.2.4 Summary

People often search for health information to fulfil an information need related to a health concern. However, the communication channels utilised by health consumers are expanding. Increased availability of the Internet has enabled rapid access to a volume of health-related content focusing on a diversity of health matters. Nevertheless, gaining entrance to this wealth of health knowledge can require the health consumer to have a level of proficiency in using electronic searching tools. Moreover, whilst various certification schemes exist to help consumers differentiate trustworthy health resources [161, 162], in many cases, the responsibility of judgment falls to the information seeker themselves. This would suggest that further analysis of consumer skills for locating and using online information could be useful.

Patient participation in choices about their health can be supported by the provision of relevant and comprehensible health education. Patient education is often enacted through the availability of medically verified written information documents. However, the usefulness and efficacy of these materials can be reduced by the inaccessibility of

the content. This may be due to the format and composition of the language and images used, however simultaneously the health literacy skills of patients may comprise an impediment to engagement with health materials. Furthermore, health literacy can affect how an individual reads and processes health information. Electronic applications have been investigated as an approach to assist less health literate individuals to engage with health information. The efficacy of many research applications in this area would suggest that further investigation of electronic tools could be beneficial. Patients are characterised by diversity with relation to information needs and health literacy skills. This suggests that a personalised approach to providing health information could assist with fulfilling the varying information needs of patients. The next section will further discuss how personalisation has been employed in digital healthcare services, and the associated methods of implementation.

2.3 Personalisation

Chronic diseases including diabetes, heart disease and respiratory diseases are the most prominent causes of death and disability globally [163]. The increasing prevalence of these conditions has been associated with health risk behaviours including unhealthy nutrition, smoking and reduced physical exercise [163, 164]. However, these behaviours are modifiable [164]. Approaches to encouraging positive behavioural change include the use of tailored health communications. *Tailoring* is a process in which information about an individual is used to adapt the composition and delivery of a health message, in order to increase the likelihood of achieving the desired outcome of the message [165, 166]. Health interventions may be tailored using factors such as demographic characteristics, psychological factors such as responsiveness to health information or willingness to change, or the risk behaviours of the user [167, 168]. Various reviews and meta-analysis of printed and electronic health behaviour interventions have found tailored interventions to be efficacious in encouraging positive behavioural change in relation to smoking cessation, physical activity and healthy nutrition [169-172]. Tailoring is often achieved by employing a combination of three strategies; *adaptation*, *feedback* and *personalisation* [165, 167, 173]. *Adaptation* involves matching the provided content to the status, needs or interests of an individual [165, 166, 173]. *Feedback* mechanisms utilise user information, collected prior to or during interactions, to provide responses which relate to the user's current behaviours, attitudes or goals [165, 173]. *Personalisation* processes focus on increasing the personal

relevance of the information for each user [165, 173]. A major focus of this research project includes development of a solution to provide personalised health information, therefore the remainder of this section will concentrate on various strategies to achieve personalisation in healthcare applications.

2.3.1 Personalised Healthcare Applications

Personalisation is a tailoring mechanism that involves incorporating recognisable attributes of an individual into the content so as to enhance the self-relevance of the information for that individual [165, 173]. It has been suggested that enhancing the personal relevance of a health message may increase attention towards the message, and motivate central processing, and heighten receptivity to the message [165, 167, 170, 173, 174]. For example, a comparison of the effectiveness of tailored and non-tailored text in health materials found that personalisation and feedback were more strongly associated with smoking cessation activities when compared to non-tailored materials [175].

Many healthcare solutions utilise personalisation to focus on the specific needs and goals of a user [176]. For example, the GlucoFit application [177] produces personalised daily and weekly exercise goals to encourage a diabetic patient to lead a more healthy and active lifestyle. Similarly, the METABO Diabetes-Lifestyle Recommender (Med-StyleR) [178] encourages long-term lifestyle changes by producing personalised nutritional and activity recommendations for a user. Patient-specific information such as daily food intake, physical activity, glucose level, age, gender and weight are utilised to recommend recipes that are suitable to each individual's specific needs and condition.

Personalising content can produce information that is more relevant or applicable to a user's behaviours, thereby increasing the usefulness of the service. For example [179] describes a website that delivered personalised feedback with the aim of reducing alcohol consumption amongst college students. Feedback was tailored as appropriate for a student's consumption patterns, associated beliefs, lifestyle and risk behaviours, and applicable website content was also highlighted. A controlled clinical trial found that female student that used the website had reduced alcohol consumption and less risk negative consequences when compared to female students that used a standardised

education website. Personalisation can also improve service provision by reducing information overload, and eliminating the need to search through large amounts of information [180, 181]. For example, [182] describes a web-based application which selects and ranks diabetes related articles from online databases in accordance with a user's profile and search terms. Thereby the volume of information provided to the user is reduced whilst the specificity of the provided information is increased.

Personalisation has also been found to foster user engagement with online applications. For example, various studies have identified increased user attention towards tailored objects included with generic content. Bang and Wojdyski [183] used eye gaze analysis to evaluate visual responses to personalised and non-personalised advertisements displayed alongside a news article. They found that participants gazed at the personalised image more frequently and for a longer amount of time when compared to the level of attention directed towards the non-personalised advertisement. This pattern of behaviour remained consistent even when users were involved in a high cognitive demand task. A similar study which examined reactions to adaptive advertising on an ecommerce based website also identified a higher rate of visual responses towards personalised advertisements when compared to attentiveness towards randomly generated advertising content [184]. Investigators have also evaluated how users engage with health information messages. Ruiter et al. [174] employed event-related potentials (ERPs) and response times to evaluate the attention resource allocation of participants when using either electronically tailored or non-tailored nutritional information. They found that increased attention resources were directed towards the tailored messages, and that the participants that read the personalised information also tended to be slower to switch attention from the tailored content when another stimulus was presented. A comparable study by these authors [185] also uncovered an early automatic preference exhibited by the participants who were reading the tailored information.

The use of tailoring in web-based health behaviour change interventions has also been associated with increased user engagement with the intervention. Danaher et al. [186] compared user interactions with a tailored, multimedia based smoking cessation application and a static text-based cessation website. Usage data indicated that the tailored application had received a significantly larger quantity of distinct visits, and

that the users of the tailored application had spent longer viewing the website content when compared to those assigned to use the static website. Strecher et al. [187] assessed whether user engagement with a tailored smoking cessation program could be influenced by including different levels of personalisation. An association was discovered between higher levels of content tailoring and increased user access to different sections of the program. Comparable participant behaviour was exhibited in a study which utilised a tailored and non-tailored intervention to encourage fruit and vegetable consumption [188]. The participants that used the tailored version of the intervention exhibited greater engagement both in terms of the breadth of content accessed and length of time spent accessing the materials.

2.3.2 Implicit and Explicit Personalisation

A prominent objective of a personalised service is to be able to predict a user's future needs in order to rapidly adapt the service [180, 189, 190]. The preference is for personalised services to adapt automatically with as little as possible direct interaction by the user [189, 191, 192]. Achieving personalisation is reliant on gathering and maintaining an accurate, comprehensive and reliable set of data about the user [191, 193, 194]. *Explicit personalisation* allows users to create and manage their own preferences set by manually entering information or configuring a pre-defined profile [180, 189, 191]. Examples of this include initialising a profile for an eCommerce or social networking site or recommender system [194, 195]. Explicit personalisation affords the user control of their profile, however, it may be cumbersome and time consuming for the user to maintain their profile and may detract them from the benefits of using the system [189, 191]. *Implicit personalisation* techniques overcome these concerns by automatically capturing information from the user while they are using a service, using this data to recognise and learn user behaviours, and tailoring the service accordingly [180, 189, 191]. By monitoring the user in different situations the user data can be dynamically updated, behaviour changes automatically detected and adaptations fed to the user. Examples of implicit personalisation include the use of purchase history to provide individual purchase recommendations in e-commerce applications [196] or the mining of social networks to suggest new associations [180]. While implicit approaches can enable quicker adaptation to behavioural changes [191] it has been suggested that inferences drawn from the data can be unreliable due to the erratic nature of user interests [197]. User data is a fundamental element of personalisation therefore

many systems will incorporate both implicit and explicit methods of collecting and analysing data. This combination approach is utilised in many major online service providers such as Amazon and Google that utilise user supplied information such as profiles and feedback, alongside implicit mechanisms [196, 198].

Recommender Systems

Issues surrounding the accumulation of user data have been a recent research focus [199]. Updating the set of information stored for a user is essential to adapting the service to new user behaviours [189, 200], and significant challenges exist in finding mechanisms to facilitate learning about the user. Recommender systems (RSs) are often used in ecommerce and social networking sites to determine which content would be most appropriate for a user's needs [196, 201]. These systems use explicit and implicit methods to collect user data and build up a model of the user's interests and preferences [192]. The user profile is then used to create personalised recommendations. The most prominently used techniques in RSs are Collaborative Filtering (CF) and Content-based Filtering [192].

CF is widely used to match users with retail items, social groups or news items [196, 198, 201, 202]. It collects user ratings and uses these to learn about the user. These ratings can be explicitly supplied by the user, for example about movies or books, or implicit indications of preference inferred from purchases [203]. Recommendations are generated from the ratings given to items by users that have similar profiles [197]. Due to its flexibility CF can be applied to a range of items and is widely used in online commercial services such as Amazon, Barnes and Noble and Netflix where there is an emphasis on providing rapid and accurate recommendations [196, 203]. CF has also been used to generate personalised recommendations for users of Google News [198]. One drawback of using CF is that it can require a substantial volume and variety of ratings to work effectively [192, 203, 204]. Additional problems faced by CF systems include the *new user problem* which occurs when the system has insufficient data about a user to generate recommendations for them [203]. Similarly the *new item problem* occurs when an item has not received enough ratings for it to be recommended to users [203]. *Reduced coverage* issues occur when user ratings relate to portions of the available items and the system is then unable to make recommendations of unrated items [203]. *Neighbour transitivity* is a situation in which associations cannot be made

between similar users if they have not rated the same items [203]. Another limitation of CF is that it can have difficulty in determining recommendations for users that have distinctive tastes and preferences [204]. While CF has a collaborative approach to generating recommendations content-based filtering is focused on matching a user's preferences with content features [205]. This approach analyses the content of favourably rated items to identify features, which are then correlated with the user's preferences. The content features are then used to identify items with similar features and these are recommended to the user. Although content-based filtering can be successful in matching a user to suitable items it also has some limitations. As the recommendations are directly associated with the user's tastes the range or variety can be quite limited [204]. In addition, the *new user problem* may arise in which a user cannot be given accurate recommendations until the system has sufficient data to learn their preferences [204]. This discussion has highlighted the strengths and limitations of both recommendation methods. *Hybrid recommender systems* attempt to overcome these issues by using both approaches to determine recommendations for a user [197, 206].

To date RSs research has largely concentrated on areas such as ecommerce, entertainment, and social media [207], and there has been less focus on the application of RSs to health informatics [208-210]. However, the use of Health Recommender Systems (HRS) is a growing research area [211]. For example, Roitman et al. [210] describe a recommendation system which aims to increase patient safety by providing tailored health information to aid medication related decisions. Web-based health content is classified using several criteria including key-words, credibility and severity. The information can then be matched to a patient's clinical status as specified by their personal health record profile. Sanchez Bocanegra et al. [212] describe a semantic-based recommender system which enriches online health information provision by adding trustworthy health content links to YouTube videos. Keywords are extracted from video metadata and matched with SNOMED-CT terms, which are then used to identify credible health education links on MedlinePlus. Other prominent areas of focus include the use of HRS to provide nutritional [213, 214] or lifestyle advice [215, 216]. Farrell et al. [215] propose a novel similarity based approach to recommending health lifestyle changes. They suggest that the uniqueness of human behaviours can limit the effectiveness of traditional CF approaches that utilise group ratings. Instead they

propose using an individual's past behaviours to extract patterns of behaviour that could be applied to create recommendations that assist with achieving health goals.

Issues Associated with Recommender Systems

Recommender systems will filter choices in accordance with the stated or inferred behaviours, characteristics or preferences of an individual. Consequently, the presented content will normally match an individual's specific needs, tastes or habits. However, there are concerns that the use of recommender algorithms may restrict the variety and range of choices available to an individual, and limit access to novel and atypical ideas or experiences [208, 217, 218].

The concept of the *echo chamber* effect (or *filter bubble* effect) refers to a situation whereby individuals are exposed to ideas or content which is closely aligned to their own beliefs, experiences or behaviours, with limited exposure to ideas or opinions that deviate or conflict with their own [217, 219]. For example, group networks formed on social media applications can act as echo chambers whereby discourse within the group largely reflects and reinforces the viewpoints of the group members [219]. In recent years echo chamber dialogues have been evident in various social media based political and scientific discussion groups [220-222]. Search engines typically personalise search results by utilising a user's browsing history, profile or geolocation [223, 224]. The main objective of this personalisation is to tailor the results to suit the user's perceived needs. In the sphere of online health information seeking there is a concern that personalisation algorithms could reduce the variety and scope of health information provided to the user [224]. Subsequently a user may interpret the provided information as confirmation of their preconceptions about a health topic, without the consideration of ancillary information or opposing viewpoints [224]. Moreover, tailoring health search results can reduce the opportunity for serendipitous discovery of novel or supplementary information related to a health topic. There are also fears that the reinforcement of erroneous or misleading health information within online echo chambers could negatively influence health decisions [224]. Research on the effects of echo chambers on public discourse is notably increasing in the areas of politics [220, 222] and environmental issues [221], however a review of the literature did not uncover any studies which specifically focused on online health information seeking.

2.3.3 User Modelling

In order for user data to support service personalisation it must be managed in a formalised structure [191]. *User modelling* is the process of establishing which user characteristics are relevant to a service and designing a structure, a *User Model*, to represent these. Kobsa [225] defined user models as “*collections of information and assumptions about individual users (as well as user groups) which are needed in the adaptation process*”. The user may be represented through a wide range of attributes including their behaviours, health status, capabilities and skills, interests and preferences [180, 226, 227]. The user’s knowledge of the service and their goal and objectives for using the system can also be modelled [228]. The user model is essential to facilitate adaptation of the service to the needs of the user, and the quality of personalisation can be affected by the accuracy, granularity and relevance of the data stored [225, 226, 229-231]. In order to preserve the accuracy of the user model it is normally stored and maintained in a separate layer from the personalisation logic [193, 232]. A *User Profile* is an instance of the user model which retains the unique attribute values for an individual user.

User modelling techniques are characterised as having a *data-driven* or a *knowledge-driven* approach [226, 233]. Data-driven approaches utilise data mining and machine learning techniques to detect user behaviour, identify user properties and produce profiles of the user [226, 234]. Large sets of user data are collected and pre-processed to extract noise and inconsistencies. Machine learning and data mining techniques are then used to identify patterns of user behaviour and produce a base model of a user [226, 234]. This is followed by validation of the user model which is usually implemented using domain knowledge, visualisation tools and can include a user model designer [226]. A wide range of machine learning algorithms can be used to identify behavioural patterns. For example [235] uses Markov Models with web server log files to learn about user interests and to predict which page they will access next. Other algorithms that can be used include Bayesian Networks and Artificial Neural Networks [226]. Although these provide effective approaches to user modelling there are a number of associated problems including the requirement for a large input dataset [226, 233] and with collaborative approaches there can be issues with determining the distinctive habits of individuals in a group [236]. Knowledge-driven approaches use expert knowledge and real world observations to construct a model of the problem domain and logic rules

[233, 237]. These approaches produce fixed models of users and match each user to the most comparable model. Reasoning engines are utilised to infer situations that can be used to shape the service delivered to the user. A prominent knowledge-based approach is the use of ontologies to formally represent a knowledge domain and support reasoning mechanisms [233]. An ontology is a model of a domain, it describes the concepts of the domain and the relationships between these concepts [238, 239]. A more detailed description of ontologies is provided in Section 2.4. In personalised services ontologies are often used to model both the user and their context.

Ontological User Models

One benefit of using ontologies is that they can be re-used in different personalisation systems. This can reduce the costs associated with developing a user model and also supports data exchange between applications that use the same ontology. A review of the literature revealed several generic user model ontologies that can be used in a range of domains. For example, the General User Modeling Ontology (GUMO) [240] is a top-level ontology for managing user and context models. It is a Web Ontology Language (OWL) ontology that represents the user in three dimensions. The first dimension, *auxiliary*, categorises the user characteristic into a generic grouping, the second dimension, *predicate*, specifies a particular aspect of this group and the third dimension, *range*, indicates a factor of user interest [240]. Incorporating this structure renders GUMO the expressiveness to represent user interests, preferences, goals and location. GUMO can be utilised in decentralised personalisation systems where it can support standardisation of user models between systems. This suggests that generic user ontologies could be used to support personalisation in different environments.

The Cold Start Problem

One of the major challenges for personalisation systems is acquiring sufficient amounts of user data to support personalisation mechanisms. The *cold start* problem normally occurs with new users for whom there is insufficient data to initialise a user model [189, 193, 241]. A similar situation can also occur with infrequent users for whom limited data has been accumulated. Solutions proposed in the literature include the sharing of user data between personalisation systems. This would enable systems to automatically initialise user models for new users and may relieve the user of the burden of having to

enter profile information. Acquiring a richer set of user data could also enable a service to expand the functionality it provides [229, 242].

A variety of approaches to user model interoperability have been proposed. *Standardisation-based user modelling* is based on the use of standardised, generic ontologies for user modelling [193, 229, 242, 243]. Each system uses a common user model ontology thereby facilitating standard representation and transfer of user data between systems. A number of ontologies are available including GUMO. Many of these can be extended to suit the particular requirements of the personalisation application. There are, however, several issues that may limit the adoption of this approach. The ontologies which are available may be too large or generic to suit the scope of a system [229] and if a new ontology is required it can be time consuming to develop and reach consensus on a new model for data exchange [193]. *Centralised approaches* use a *user modelling server* to store and maintain a user model which can be accessed by distributed applications [193, 229]. Servers can provide advanced features for management, querying and security of user data and can also provide larger collections of user data than would be gathered by stand-alone systems [229, 244]. Additional advantages of having a central repository of user data include the ability to check the consistency and coherence of information gathered by the different applications [232]. This approach does, however, have several disadvantages. Having a centralised model requires an application to adapt to the central data model [193] and moreover a central store of data is a single point of failure [229]. This risk can be mitigated by having mirror servers [229].

2.3.4 Context

If a user moves to a different environment they may encounter factors that will have an effect on personalised service provision. Alternatively, if changes are introduced into the environment of a static user this may also impact the level of personalisation available. Contextualisation, “*the capability to offer the right information and services, tailoring them to the right device, to the right user, at the right time and location*” [195] is a process for managing contextual factors and using these to adapt service provision accordingly. The use of contextualisation in personalisation systems is an evolving research area [199, 245, 246]. To understand how context information could be used in personalisation systems we firstly need to identify the features of a user’s environment

that constitute their context and determine how a user's needs change in response to changes in these features.

A review of the literature indicated that conceptualisations of context are wide ranging and include changes associated with the user, or factors that are external to the user including location and environment. Indeed there is no definitive consensus of what constitutes context [189]. The definition of context proposed by Dey [247] is “*any information that can be used to characterise the situation of an entity*”, which is, “*a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.*” Many definitions of context include spatial and environmental elements such as time, location, weather, temperature, noise level and lighting [248-250]. Changes in these factors can obviously alter a person's objective in using a personalised service. Context awareness can be devised from localisation mechanisms, mobile device characteristics, network availability, environmental factors, mobility patterns and transportation use [195, 251, 252]. This information can be used to adapt the service to the preferences of the user. A task is described as an activity that the user completes to achieve a specific goal [248]. Many definitions of context place significance on the task that the user is carrying out when using a service [248, 250]. For example, [253] presents a smart-phone application that provides context-based reminders for persons with memory impairments. This application enables a user to enter tasks and evaluates contextual factors such as time and location so that a user may be prompted to carry out a task within a certain time period or whenever they are close to a location convenient for fulfilling a task.

Modelling Context

The aim of a context-aware service is to detect and gather information about changes in a user's context and dynamically adapt the service provided without any user intervention [246, 254, 255]. A system's capability to provide a contextualised service is determined by the type of context information that is being monitored and how it is modelled [191, 254]. There are many challenges associated with collecting and modelling context data. Context information can be collected from heterogeneous sources and can differ in accuracy, completeness and format [233, 245]. Data modelling approaches must therefore account for uncertainties and deficiencies in the data. A

context data model should have the flexibility to represent different types and granularity of context data and inter-relationships [245, 250, 256].

A *Context Model* is a formal, structural model of context data [254]. Various techniques for modelling context data can be used, including object-role based and ontology-based models. Object-role models produce a context model in graphical notation [255, 256]. This model captures context as classes, facts and properties and relationships as edges in between. A strength of this approach is that it presents an understandable model of context for the developer and can be used at different stages of the software development cycle [245, 256]. A suggested limitation is that it does not differentiate between different types of context data [245]. Ontologies are perceived as a formalised, expressive and reusable means of capturing context information [233, 246, 255-257]. Furthermore, using ontologies to represent context data facilitates the use of reasoning to ensure the correctness of the model and deduce new knowledge from the model.

2.3.5 Summary

The use of information technology to deliver healthcare solutions has the potential to expand the availability, accessibility, sophistication and cost-effectiveness of tailored communications [167-169]. Many electronic healthcare tools incorporate personalisation and contextualisation with the aim of motivating patient engagement, and several reviews have suggested that personalised health interventions could be a viable approach for promoting positive health behaviours. However, while the efficacy of personalised interventions has been established there is a shifting research focus towards considering which particular conditions of personalisation are most effectual and why this is so [166, 170, 171, 175]. Proposed research directions include analysis of the effectiveness of different tailoring mechanisms to achieve particular goals [171]. This would suggest that challenges still remain in the evaluation of personalisation systems.

Personalisation is facilitated by user information and a key challenge is to find ways to gather and model information about the user. A widely used solution is the implementation of user models, uniform software structures that hold user attributes and the relationships between these. Approaches to user modelling can include a data-driven or a knowledge-based approach. Data-driven approaches concentrate on data mining

and machine learning approaches to identify user characteristics while knowledge-driven approaches utilise domain expert knowledge to build models of the user [193, 226]. Ontologies provide expressivity to capture a range of user data and establish constraints on relationships between data. This facilitates the capture of an accurate and reliable model of the user and the use of reasoning technologies to infer new knowledge about the user. The collection of context data is also vital to identifying changes in the user's situation and adapting the service to suit their new requirements. Semantic technologies can be used to infer associations between context and user data. This research project concentrates on the knowledge-driven approach and the use of ontologies to develop user models. The next section will focus on ontologies and describe how these can be used to capture knowledge in health domains. The application of semantic web technologies to support personalisation processes in healthcare applications will also be investigated.

2.4 Ontologies

The use of ontologies and rule languages to support personalisation mechanisms is one of the key focuses of this research project. As previously discussed, ontologies are a prevailing approach to user modelling [193, 243], and are also widely used in context modelling [245, 254, 256]. Corcho et al. [258] defined ontologies as “*formal, explicit specifications of shared conceptualizations*”. An ontology comprises a world view of a domain that is conceptualised through an arrangement of concepts, specifications (definitions) and inter-relationships, and shared agreement on the meaning of these components [256, 259].

An ontology is represented as a vocabulary and is encoded in a formal ontology language [256, 259, 260]. Within the ontology the domain concepts are represented as *classes* [261, 262]. Domain concepts may be perceived as entities within the domain that are significant to the application. Classes are arranged in a taxonomy of superclass-subclass relationships which are referred to as *subsumption relationships* [262]. A class may have *properties* that define its relationships with other components of the ontology. Object properties denote relationships between classes whilst datatype properties denote attributes or data values of a class [262]. Properties can also be used to define restrictions on a class, for example a property may define the cardinality of a relationship between two classes. An instantiation of a class which contains data and

object values is referred to as an *instance* [261]. Fig. 2.1 illustrates a class hierarchy containing classes and relationships. This is a sub-tree of an ontology that models information about diabetes medications. As can be perceived a subclass is a more specialised definition of a superclass.

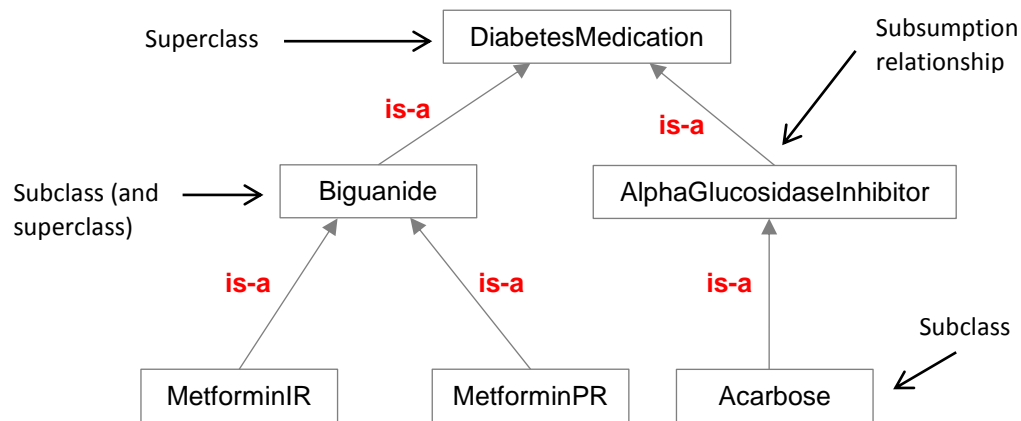


Fig. 2.1. Taxonomical model of ontology sub-tree which represents diabetes medications.

2.4.1 Ontology Representation Languages & Reasoning Technologies

Personalisation, which can be recognised in many Internet-based services, is facilitated by the abundance of online information that is available about the user. Data can be gathered from user contributions such as reviews or ratings and can then be correlated to a user's interests or preferences. Online activities such as purchasing items, booking tickets and building a network of friends on a social networking site can also be used to deduce knowledge about the user. There are, however, challenges to utilising this wealth of data. Whilst computational algorithms can retrieve and organise information, the meaning of information can only be deduced by humans. Information processing would produce more usable results if computers were able to process semantics and analyse information in the same way humans do. The *Semantic Web* proposes a solution to this problem through technologies that associate information with semantics [263-266]. The Semantic Web is a W3C architectural plan for the web, that will enable machines to process information in correspondence to its (human assigned) meaning [264, 265]. The underpinning principles of the Semantic Web were described in a 2001 article by Berners-Lee et al. [263]. A primary theme involves the reformulation of online data sets

so as to facilitate semantic interpretation. Much of the content of Web 2.0 exists in a variety of structured, unstructured and semi-structured formats, which can prove challenging for automated searching technologies. However, implementing semantic web based solutions requires data to be available in structured formats that will add meaning to information, and thereby facilitate automated reasoning on the data by inference technologies. Ontologies are regarded as an underpinning component of the semantic web as they provide a standardised means to represent the structure of knowledge and define the semantics of this knowledge [263, 267]. Moreover, the application of inference technologies enables reasoning on the ontology, and the deduction of novel information [266, 267]. By adding meaning to data, ontologies also enable the integration of heterogeneous data sets, and facilitate interoperability between different systems [268-270]. Moreover ontologies also have the potential to enhance the accuracy of web searches [263]. There have been a number of W3C recommendations for human and machine understandable languages that can be used to construct ontologies. These include Resource Description Framework (RDF) [271], RDF Schema (RDFS) [266] and Web Ontology Language (OWL) [239].

RDF is a language which uses Extensible Mark-Up Language (XML) syntax to make assertions about web resources. It is domain independent and allows developers to describe resources using their preferred vocabulary [266]. Resources are identified using URIs and are described through a data model called a triple. In a triple a resource is the subject, and is associated with a predicate which has an object. Each RDF triple can be represented as a *directed graph* or encoded as a document *statement*. One of the primary purposes of RDF is to assist data exchange [265] and this is facilitated by the use of RDFS. RDFS adds meaning to RDF by describing how the vocabulary should be interpreted [266]. In this way it enables machines to share RDF information by mapping from one schema to another. It provides a more complete vocabulary when compared to RDF and can define classes, properties and class hierarchies. An endeavour by W3C working groups to provide more flexibility and expressiveness to describe web resources resulted in the development of OWL. OWL is layered on top of XML, RDF and RDFS and uses XML style syntax, whilst also having a greater range of features to capture complex information. In addition to defining class hierarchies and object and data type properties, OWL also allows class constraints to be defined. These include cardinality restrictions, transitive and inverse properties, class disjointness and class

equivalence [262, 266]. The original OWL standard was superseded in 2009 by the OWL 2 specification [272]. OWL 2 is available in two sublanguages. *OWL 2 Full* is an extension of RDFS and uses the full expressiveness of the language. However, as it is interpreted under RDF-Based Semantics it does not have full reasoning support. *OWL 2 DL* is a subset of OWL 2 Full and provides a more limited use of the OWL constructs, but in contrast to OWL 2 Full it is decidable. OWL 2 includes three *profiles*. The *OWL 2 EL* profile [273] was designed to facilitate efficient reasoning on larger ontologies (such as biohealth ontologies) which comprise a large quantity of class and property axioms. The *OWL 2 QL* profile [273, 274] provides the expressiveness to represent conceptual models including UML class diagrams and ER diagrams. It facilitates querying of a relational database through an ontology. The *OWL 2 RL* profile [274] provides a syntactic subset of OWL 2 that can be implemented using rules. This profile affords scalable reasoning capabilities alongside significant expressivity. Querying support for RDF based information is provided through the SPARQL Protocol and RDF Query Language (SPARQL) [275]. SPARQL implements queries as graph patterns which execute on RDF triple stores, and also provides functions for advanced querying expressions including result aggregation. In addition to the realisation of several semantic web standards, a notable collection of software tools have been created to support ontology authoring, visualisation and manipulation. A detailed list of tools is maintained by the W3C at [276].

Expressing an ontology in a formal language such as RDF or OWL facilitates the use of a software reasoning tool to check the correctness and consistency of the model [266, 277]. For example a reasoner can be used to find duplicate instances and inconsistent classes whose conditions determine that it can never have any instances [265, 277]. A reasoner can also be used to infer new knowledge from the ontology by discovering new relationships between classes [278]. *Classifying* an ontology involves checking whether a class could have more than one parent class and produces an *inferred* class hierarchy of all possible subsumption relationships [262, 265]. This may be helpful for engineering and maintaining large generic ontologies [262]. In addition, a reasoner can identify relationships between disparate concepts [279].

Usage of Semantic Web Ontologies

The use of ontologies in various scientific and biomedical research fields is regarded as one of the major successes of the semantic web movement [267]. The bioinformatics research community in particular has been characterised by significant interest in ontology-based solutions [280]. This research domain is characterised by a need to integrate high volumes of data from disparate sources and support high throughput scientific methodologies [281]. Ontologies have provided a means to define common vocabularies for distinct data sources, enabling data sharing and reuse, and moreover are frequently employed as a means for annotating data and resources, and supporting knowledge management for decision support systems [268, 281-284]. Several large biomedical domain ontologies are also available including the Gene Ontology (GO), the International Classification of Diseases (ICD), and the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT) [268, 270, 281, 282]. The proliferation of large domain ontologies alongside smaller specialised ontologies has led to several projects which promote ontology reuse. The Unified Medical Language System (UMLS) [285] has facilitated the merging of existing ontologies in order to alleviate the problems associated with ambiguous representation of domain concepts across comparable ontologies. In contrast, the Open Biological and Biomedical Ontology (OBO) Foundry [286] provides shared principles for ontology development and reformulation, and encourages reuse of existing ontological components.

Other areas of scientific research are recognising the benefits of using semantic web technologies as a means to support data discovery and increase the accessibility of scientific data. Narock and Wimmer [287] noted growing adoption of semantic web technologies in the geosciences (astronomy, earth science, hydrology, and solar-terrestrial physics), with a particular focus on the development, usage and validation of ontologies and frameworks. Other application areas of semantic web technologies include industrial activities. Critiques of manufacturing systems have identified the use of ontologies, reasoners and semantic web frameworks, to capture expert knowledge, facilitate information reuse, connect heterogeneous systems and reconfigure manufacturing processes [288-290]. Ontologies and semantic web services have also been applied to support knowledge management, integration and querying in various enterprise applications [291]. Interest in semantic web solutions is also expanding in the commercial and media sectors. For example, a metadata based infrastructure (including

the use of ontologies, RDFa and SPARQL querying) was created to assist content management and publishing on the BBC sports site for the 2010 World Cup and 2012 Olympics [292, 293]. The electronic retailer Best Buy employed the eCommerce ontology *GoodRelations* [294] to publish metadata about their website, with the aim of increasing online visibility [295]. This resulted in a 30% increase of search traffic [295].

This review provides a short critique of the expanding use of semantic web based ontologies in a range of application areas. However, there are some drawbacks to using RDF or OWL based ontologies to express application or domain knowledge. These will be discussed in the next section.

Issues Associated with OWL DL Ontologies

OWL DL is based on a description logic (DL), and provides both a formal semantics for expressing knowledge, as well as decidability. However, it has been suggested that the expressive power of DL based languages can be inadequate to accurately represent certain categories or characteristics of biomedical knowledge [296-298]. For instance, OWL DL object properties are binary relationships which link only two individuals. However, certain cases of biomedical knowledge can require ternary or higher degree relationships to be represented [298]. For example, certain biological entities may exist in different states at different times and thereby require that a third, temporal dimension of an assertion be recorded [296, 297]. An ontological model of a domain is defined as being an explicit and unambiguous assertion of knowledge within a domain. However the domain of biology is characterised by inherent uncertainties regarding phenomena, and exceptional or probabilistic events [298]. Challenges exist in representing uncertainties and presumptive knowledge in DL based languages. Schulz et al. [297] refer to the existence of *background knowledge*, “*statements that are assumed to be at least typically (but not necessarily universally) true*” for any domain. For instance, one category of background knowledge is *default knowledge*. Default knowledge about a biomedical entity relates to canonical information that is expected to be true but can be deviated from in some cases [297, 299]. In general humans are able to understand default knowledge by making implicit assumptions about exceptional cases. For instance, [297] uses the example of anatomical canonical information which defines that a hand typically has a thumb. Nevertheless, having default knowledge of anatomy denotes that this is not always the case. However, OWL DL class definitions assert

conditions that are true for all class instances in all cases (e.g. an individual of class hand always has a relationship with an individual of class thumb) [298], and do not provide the expressiveness to accurately capture atypical or exceptional conditions that may occur [297].

Another category of information that can be difficult to capture in an OWL DL ontology are *dispositions* [297]. Dispositions are statements regarding events or conditions which have a tendency to occur for an entity under certain conditions, however the entity can also exist without the disposition ever being realised [296, 297]. For example, antibiotic medications can cause an allergic reaction in patients under particular circumstances, however this outcome may never occur. Representing dispositions in an OWL DL ontology can require complex statements which are computationally expensive. Consequently this can affect the scalability of the ontology and can impact the efficiency of reasoning tools on the ontology [296, 297]. It has been suggested that using Ontology Design Patterns (ODPs) can overcome some of the limitations of OWL DL with relation to modelling object relationships that involve more than two individuals, and capturing knowledge related to exceptional situations that can occur with biological entities [298].

2.4.2 Ontology Engineering

Biomedical and healthcare ontologies may be developed through the re-use of established vocabularies or through the construction of an original ontological model. Ontology engineering has been defined as “*the set of activities that concern the ontology development process, the ontology lifecycle, the principles, methods and methodologies for building ontologies, and the tool suites and languages that support them*” [258]. An ontology is typically developed through a series of activities during which an understanding of the domain is achieved, followed by phases of specification, conceptualisation, formalisation, implementation, evaluation and maintenance [300]. Many of the activities are iterative whereby previous stages are revisited as the ontology is continuously revised and refined. Several formal methodologies have been devised which define the stages of development and the tasks to be undertaken at each stage [261, 300, 301]. Formal methodologies have been extensively applied to guide the engineering of biomedical and health domain ontologies. For example, the *101 method* [261] was employed during the development of ontologies that focused on heart rate

turbulence [302], the enactment of clinical guidelines within a medical organisation [303], and antimicrobial-microorganism [304].

Ontology development is a knowledge focused process which involves ontology engineers and domain experts. However engineering an ontology can be time consuming, labour intensive, and error prone [305-307]. This has led to an increasing research focus on developing more efficient methods and tools that will support collaborative ontology engineering activities. Collaborative processes are being increasingly used for the development of large ontologies where engineering would be unfeasible for a small team [308]. For example, the development of large biomedical ontologies is facilitated by increasing contributions from scientists [308, 309]. Much of the literature in this area addresses the implementation of collaborative tools and methodologies and the practices and skills of contributors [310, 311]. For example, a study to evaluate the social dynamics behind collaborative processes is described in [312]. The authors propose that by analysing quantitative data from collaborative projects insights can be made into the social, lexical and behavioural facets of this approach. Social aspects can include whether contributors work alone or in collaboration. Lexical aspects can relate to stabilisation of the ontology, for example whether the size of the vocabulary is constant or continues to grow. Behavioural facets can relate to whether higher level concepts are defined first (a top-down approach) or development begins with more granular concepts (a bottom-up approach). The authors suggest that understanding characteristics such as these can contribute to ontology evaluation and thereby can help to improve the quality of ontologies.

2.4.3 Biomedical and Health Ontologies

Ontologies provide the expressiveness to represent a diverse range of clinical data and the associated interconnections between data. The benefits of using ontological modelling to specify medical concepts are highlighted by Kökciyan et al. [313]. They addressed the issues associated with ambiguity and inconsistency in radiology reports by developing an ontology (ONLIRA) to describe liver imaging observations. Moreover, they also detail how ontological features such as classes, properties and constraints enabled a thorough and precise model of the liver, hepatic veins, liver lesions and associated relationships to be constructed.

Ontologies have been widely utilised as a means of standardising medical terminologies and providing validated semantics for medical and health based data [268]. For example, one of the most prominently known healthcare terminologies is the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT) [314], a multilingual clinical vocabulary that can capture a range of clinical information, and also provides the core terms used for the electronic health record. Large terminologies can be re-used and extended as appropriate for individual medical applications. For example, Soguero-Ruiz et al. [302] used SNOMED CT to represent a model of heart rate turbulence that could be included in the EHR. They mapped clinical, anatomical, electrophysiological and pharmacological features of the HRT domain to SNOMED CT concepts and extended the representation by adding 19 new concepts. Constructing the HRT ontology within the SNOMED CT model provided a standardised and accessible tool that could be applied to support decision making in cardiovascular risk stratification. Biomedical ontologies can also be used to provide a common vocabulary for different knowledge standards, and are often used to facilitate data sharing between heterogeneous medical systems. For example, Ryan [315] describes how mapping between SNOMED and the Health Level 7 (HL7) standard could underpin data exchange by coding HL7 messages using SNOMED. Furthermore, a solution that uses HL7 and SNOMED for the reuse of Clinical Decision Support Systems (CDSS) in different Health Information Systems is described in [316].

A reasoning component can utilise an ontology and rules to make inferences about associations in the data [256]. Semantic rules may be specified using semantic web rule languages such as Semantic Web Rule Language (SWRL) [317] and SPARQL Protocol and RDF Query Language (SPARQL) [275]. Some examples of systems that combine the use of ontology and rules are the following. Bright et al. [304] constructed an ontological model of the domain of antimicrobial prescribing. SWRL rules were developed to infer knowledge from the ontology, which was then used to derive alert messages that indicated an incompatible scenario for antimicrobial prescribing. Contextual information is often used in healthcare systems to support the fluctuating needs of a patient in different environments. This is especially important in home-based telemonitoring scenarios where patient specific information must be supplemented with contextual information in order for a physician to make informed decisions [318]. The development of an ontology that supports a telemonitoring service for patients with

chronic conditions is described by Lasierra et al. [318]. A Monitoring Task class defines the measurements to be collected by a patient including vital sign measurements and information regarding the context in which a measurement was taken. SPARQL rules are used to determine the sequencing of tasks whereby a monitoring task may trigger a specific analysis task which evaluates whether a measurement is within a specific range. Semantic reasoning can also facilitate personalised supervision for each patient as a physician can define distinct observation parameters for an individual and thereby tailor the service to the health needs of that patient. An ontology-based architecture that uses contextual information to personalise non-pharmacological interventions for people with dementia (PwD) is presented by Navarro et al. [319]. The ontology captures the knowledge required to identify behavioural and psychological symptoms of dementia (BPSD). This includes aspects of the PwD such as their health and cognitive status, personal characteristics, and the physical and social context in which a BPSD occurred. SWRL rules are utilised to reason on this information and match a BPSD with an intervention that is suitable for the context and characteristics of the PwD.

2.4.4 Knowledge-Driven Approaches to Ehealth Literacy Applications

Ontological knowledge models have also been used in applications that support access to health information for persons with varying levels of eHealth literacy. Gradie et al. [320] present *SciReader*, a web-application that aids reading comprehension of medical and scientific documents by enabling readers to automatically view definitions of unfamiliar words in the text. The vocabulary employed in the application is derived from several ontological and dictionary sources including the Gene Ontology, Open Biomedical Ontologies, WordNet and the NCI Thesaurus. The authors propose that *SciReader* can enhance medical literacy by augmenting understanding of the health information for the user, and reducing reading time. Demelo et al. [321] focused on the difficulties that users can face in articulating their health information needs. They developed Ontology-Driven Visual Search and Triage Interface for MEDLINE (OVERT-MED), a browser-based tool which automatically responds to search term entry by suggesting related health search queries. The queries include terminology derived from the Human Phenotype Ontology (HPO), and the interface also provides an option to link to the HPO web browser to find definitions for unrecognised terms. It is suggested that use of this tool can assist users with augmenting their knowledge and vocabulary related to a specific health domain. Several applications have attempted to

tailor online health information as appropriate for the health status and capabilities of the user. Al-Busaidi et al. [322] propose an ontology-based architecture that will provide online health information which corresponds to a patient's health information needs, as defined in their patient medical record. An ontology is utilised to represent patient medical record data and associate this with terminology used to describe online health information resources. The ontology is then utilised to expand online searches for health information by supplementing a user's search queries with synonymous health terms. The authors denote that this mechanism may provide a user with additional useful information which has not been specified by their original search criteria. Rohrer et al. [323] propose an ontology-based recommender system which uses website quality characteristics to match online health content to a user's profile and information searching goals. The system includes an ontology network which comprises a collection of ontologies which represent pertinent factors including health domains, web pages, quality factors and user profiles. The recommendation process evaluates web page quality dimensions, including readability and timeliness, so as to match web resources to a user's personal characteristics and the context in which they are searching for information.

2.4.5 Summary

Ontologies provide a human and machine readable method for representing a model of the knowledge in a domain, and have been widely utilised in the biomedical and healthcare domains to represent data semantics for a wide range of clinical and contextual data. Semantic web technologies provide a means not only to create an ontology but also to reason on the ontology and facilitate personalisation processes. As discussed, ontology-based architectures have facilitated healthcare provision for a range of health conditions and have also enabled interoperability between different clinical systems. This would suggest that semantic web based technologies could be applied to novel healthcare scenarios that focus on personalised healthcare provision. The use of an ontology to capture the domain of health education will be further investigated in Chapter 4 of this thesis.

Developing an ontology can be a complex and time consuming undertaking. However, several methodologies have been defined to guide engineering processes. Nevertheless, larger engineering projects may involve a quantity of specialists that may also be

geographically dispersed. The breadth of interest in collaborative approaches suggests that research is focused on finding new methods to facilitate ontology development. These approaches aim to reduce the costs associated with ontology development and enhance the accuracy of ontological models. This would suggest that research knowledge contributions could focus on identifying practical collaborative methods to assist with ontology engineering processes.

2.5 Conclusion

This chapter provided a review of the literature associated with the three main focus areas of this research project. Section 2.2 concentrated on health information seeking and described two channels through which health information can be consumed; printed patient education materials and Internet based resources. The Internet provides a convenient method of accessing health information for many consumers, however, access to this information may be limited for individuals that do not have sufficient technical searching skills. Moreover, due to the largely unregulated nature of the information available, it is also vital that health consumers have the skills to scrutinise and differentiate between the quality of different online resources. To date observational studies have evaluated the online skills of health consumers, however less studies have appraised the associations between health literacy skills and online health information seeking. Chapter 3 addresses this knowledge gap by describing a study of online health information seeking that investigates relationships with both health literacy and eHealth literacy skills.

Although generic patient education materials provide medically verified information, they may not always fulfil the personal information needs and literacy skills of all patients. A solution to alleviate these issues may be enabled through the concept of personalisation, a strategy that has been used to shape the provision of healthcare applications to the needs, preferences and capabilities of individual patients. Section 2.3 provided examples of how personalisation has been realised in various healthcare applications. However, providing a personalised health service requires a range of knowledge related to the health domain and the target user. Sections 2.3 and 2.4 provided an analysis of the benefits of using ontologies to capture clinical and health data. Furthermore, semantic web technologies provide a means by which to develop a framework which incorporates a knowledge base and personalisation mechanisms.

Chapters 4, 5 and 6 address the use of personalisation in health education provision and describe a technical framework to achieve personalised education, and an evaluation of patient engagement with personalised education.

Chapter 3

Computing Health and eHealth Literacies Using Existing Instruments and Browser-based Software for Tracking Online Health Information Seeking Behaviour

3.1 Introduction

Health information seeking is often differentiated from passive acquisition of health knowledge and has been described as undertaking purposeful activities to obtain health information, often to fulfil specific goals [122, 324]. Moreover it is being increasingly recognised that motivation to engage with health information seeking may be influenced by the personal characteristics and aptitude of the information seeker, and the context in which the search takes place [121, 122]. Whilst some patients prefer to receive health information directly from their physician or to use printed content such as patient information leaflets [14, 21, 24, 116], increased access to the Internet has introduced many health consumers to new sources of health knowledge. Many government and health organisations have embraced this new channel of healthcare delivery and operate web portals that provide scientifically based health information at the point of need. However, the open nature of the World Wide Web has also facilitated the publication of a volume of unregulated health content which can be easily reached through use of a Web browser. With increasing use of the Web to obtain health information [22, 23], it is important to understand the strategies employed by information seekers when searching online for health information. This includes the methods and tools used to search for information, the types of health information gathered and whether the information can be applied to fulfil a health need. Thereby the first research question of this project is proposed, which focuses on the online health information seeking behaviour (OHISB) of health information seekers.

- Do any relationships exist between an individual's level of health literacy and eHealth literacy, and the behaviours exhibited when seeking and using online health information?

This chapter describes a study which was designed to investigate the research question. The aim of the study was to explore the associations between health literacy and eHealth literacy, and an individual's skills to find and apply online health information to a popular health topic. As denoted in Chapter 2, health literacy is an individual's

competence to obtain, understand and apply health information [30]. The concept of eHealth literacy (also called *digital health literacy*) relates specifically to online health information seeking, and is defined as “*the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem*” [325]. Being eHealth literate is associated with advanced skills for online health information searching. Whilst much research has focused on the relationship between eHealth literacy and online health information seeking, fewer studies have investigated both health literacy and eHealth literacy as determinants. It was decided to investigate OHISB from the dual perspective of health literacy and eHealth literacy as online information gathering requires multiple interrelated skills. The online health information seeker requires competencies to effectively utilise online searching applications and filter search results. Moreover literacy, numeracy, decision making and reasoning skills may be needed to critically evaluate the information found. These skills can be effectively measured with health literacy and eHealth literacy instruments, enabling further analysis of OHISB.

The remainder of the chapter will describe the methodology, results and conclusions drawn from the study. Section 3.2 further establishes the context of the research by discussing related work in the fields of health literacy and eHealth literacy, and introduces the relationship between health information seeking and uncertainty management. The methodology by which health literacy, eHealth literacy and OHISB were evaluated is described in Section 3.3.

3.2 Related Work

3.2.1 Online Health Information Seeking and Uncertainty Management

Health information seeking is often initiated in response to a health problem or health-related anxiety [88, 326, 327]. An information need can arise when a patient’s level of knowledge is incongruent with that necessary to fulfil their health need. Uncertainty is a prevalent characteristic of health experiences such as illness [328, 329], and can occur “*when details of situations are ambiguous, complex, unpredictable, or probabilistic*” [330]. For patients and caregivers medical sources of anxiety can include the diagnosis, treatment or trajectory of a health condition, or aspects of medical decision making [331-334]. Moreover, insufficient information, or the provision of information that is

inaccurate, inconsistent, ambiguous or excessive can also contribute to uncertainty and confusion [92, 329-331, 333, 335]. One strategy by which individuals may attempt to manage their uncertainty is by actively seeking information [326, 327, 333, 335, 336]. Online health resources have been used as a means to cope with or reduce health-related uncertainty [92, 327, 332, 335, 337], and *Uncertainty Management Theory* (UMT) [328], a prominent communication uncertainty framework, has been applied to appraise the associations between online health information seeking and uncertainty management [332, 338-340]. A central tenet of UMT proposes that uncertainty is not necessarily a negative or positive experience, but that an individual will appraise the meaning of uncertainty, and the resulting emotional response will determine whether the uncertainty is evaluated as negative, positive or neutral. The uncertainty evaluation will influence an individual's behaviours in managing their uncertainty. For example, individuals for whom uncertainty is an undesirable or negative state may seek health information to augment their knowledge and thereby lessen their state of uncertainty [327, 336]. In contrast, others who perceive their uncertainty as rendering hope or optimism may strive to maintain or increase the state of uncertainty by gathering discordant information or by intentionally avoiding information that could potentially cause distress or discomfort [327, 341]. It has been suggested that social support with gathering, examining and evaluating information, for example by friends or family members, can assist with uncertainty management [342]. Web-based resources such as online communities may provide a platform for interpersonal communications which can be utilised to cope with health-related uncertainty. UMT was utilised to analyse the behaviours exhibited by the members of an online health support community for the parents and caregivers of children with clubfoot [332]. Analysis of the information exchanges revealed that the behaviours of both information seekers and information providers were focused on uncertainty management; whilst caregivers sought to acquire information that would help them to cope with the uncertainties associated with their children's illness, those that provided advice also aimed to identify and provide knowledge that would assist information seekers with managing their uncertainties.

Online searching tools can facilitate access to a diversity of health information resources that address both prominent and less known health issues. Rains [338] suggested that OHISB such as query construction and information source selection enabled the information seeker to influence and manage information acquisition processes.

Moreover, the range of information resources available provided information seekers with opportunities to identify information specific to their health needs which could subsequently be used to manage their uncertainties. A UMT based analysis of information seeking found that individuals who utilised Web-based health information resources were more successful in achieving a desired level of uncertainty when compared to individuals whom did not seek health information [338]. Researchers have also examined the associations between the particular categories of medical content available online and uncertainty management. An investigation of online searches for skin cancer information discovered that scrutinising web content related to the detection, prevention and treatment of skin cancer was not predictive of reaching a desirable level of uncertainty [340]. However, reading information that focused on skin cancer susceptibility and severity was associated with successful uncertainty management. UMT has also been utilised to evaluate how the searching behaviours and information-processing activities of online information seekers may affect uncertainty management. Rains and Tukachinsky [340] found that participants who spent longer reading health web pages, and thus were highly invested in systematic processing of the information, achieved greater levels of uncertainty discrepancy reduction. An appraisal of online searching behaviours discovered that participants who exhibited higher levels of positivity in uncertainty appraisals exhibited a more focused approach to online health information searches, visiting fewer web pages and having longer visit durations than those with more negative appraisals [339]. The investigators proposed that those with a more optimistic perception of uncertainty may be driven to select and scrutinise information sources that confirm their current knowledge, and thereby can assist with maintaining their uncertainty levels.

Health consumers are often faced with challenging and uncertain health situations and may employ various strategies to reduce or maintain ambiguity about a health-related condition. The reviewed literature suggests that the Internet may provide an effective and valuable health information channel for individuals that wish to utilise information strategies in order to manage health-related uncertainty.

3.2.2 Health Literacy and Ehealth Literacy

The increasing utilisation of eHealth applications provides an opportunity for widespread dissemination of reliable and timely health information to health

consumers. Moreover, in many cases eHealth communications can facilitate the tailoring of health messages to the particular needs and behaviours of health consumers [343]. The potential for eHealth interventions to encourage positive behavioural outcomes has been established [344, 345]. However, there are various impediments to the adoption of eHealth technologies. Ehealth tools are required to be accessible, easy to use and engaging for the audience [343]. Moreover, the competencies of the intended audience are also of significance; if consumers do not have the necessary skills to utilise eHealth tools then as a result their effectiveness will be limited [325]. The competencies to effectively engage with eHealth are reflected in Norman and Skinner's model of eHealth literacy [325]. This model comprises six core literacies that are subdivided into two main categories, (1) analytic and (2) context-specific. *Analytic* skills include the *traditional literacy* skills of reading, writing and numeracy, and *information literacy*, an individual's ability to comprehend how information is organised, and awareness of where and how to find information. Also included in this category is *media literacy*, an ability to understand the composition, context and influences of media based information. *Context-specific* skills include *health literacy* and *computer literacy*, which is the aptitude to access and use computers to resolve issues, and adapt to new software and technologies. The final literacy in this category is *scientific literacy*, which refers to an understanding of the scientific methodologies involved in knowledge discovery. The literacies combine to support the consumer's interaction with eHealth. However, eHealth literacy is context dependent and may be influenced by the motivation for seeking information, and the health status, presenting health issues and educational background of an individual. Moreover, it is possible that a person's eHealth literacy skills may advance through appropriate training. For example an intervention for HIV-positive adults achieved quantified improvement in the participants' eHealth literacy scores and skills to access Internet based health information [346]. Being eHealth literate has been associated with increased health knowledge and activation, for example through discussions with healthcare providers and participation in preventative screening practises [95, 347, 348].

3.2.2.1 Online Health Information Seeking Behaviour

Much research has identified an association between eHealth literacy and an individual's motivation to engage with online health information seeking. It has been found that those with higher levels of eHealth literacy have a higher perception of the

usefulness of Internet based health resources for making health decisions [349, 350]. Moreover, those with high levels of eHealth literacy have been identified as more frequent health information seekers [351, 352], and more likely to scrutinise the accuracy and reliability of an information source, and have a greater range of online searching strategies when compared with a lower eHealth literacy group [351]. Lower levels of educational attainment are often associated with reduced levels of health literacy, eHealth literacy and ineffective online searching strategies. One group of individuals that can have increased motivation to seek health information are parents that have health concerns for their children. Studies that involved the parents of children with life-threatening illnesses [353] and special health care needs [354] both revealed an association between not having a high school diploma and reduced eHealth literacy. Moreover, some of the parents expressed difficulties distinguishing between high and low quality information sources and sometimes lacked the confidence to act on health information [354]. However, it has also been suggested that difficulties with locating and assessing health information are not limited to those with lower levels of education. Surveys related to the eHealth skills of undergraduate nurses found that a considerable proportion of the students were unconfident in their ability to discriminate between high and low quality health resources, and to apply the information found to making a health decision [349, 350]. A study of female college students attending a health professional degree program found variation in the students' perceived searching abilities [355]. Whilst some students reported using multiple health resources, others reported that they relied more heavily on a search engine to locate health information resources. Moreover, some of the students reported limitations in search query formulation. A survey of university students attending health courses found comparable difficulties with information searching activities [356]. Although students were aware that using a single word query was not an optimal searching approach the majority of students relied on this strategy as they were unaware of how to add multiple search categories or Boolean operators. It has been suggested that information searching incompetency amongst students may limit their ability to benefit from online health information resources and may hinder them in making health decisions [356].

Another factor that is frequently associated with decreased levels of eHealth literacy is increased age. A study which concentrated on adults with chronic health conditions found that older adults and those with lower levels of eHealth literacy were prone to

navigational needs, experiencing difficulties in finding online health information, and being less assured in their searching abilities [95]. An examination of Internet skills also found that older adults sometimes experienced problems when completing tasks that called upon operational and formal Internet skills [357]. This included difficulties identify and using the browser address bar, and understanding orientation within a website. A survey that addressed the health information seeking behaviours of baby boomers and older adults found that an increase in age did correspond with a decrease in eHEALS scores [358]. However, in contrast with other research studies, the authors found that the respondents were largely positive regarding their ability to find and use Internet based health resources, although there was less confidence in their ability to differentiate between high and low quality resources. Similar results were reflected in a survey involving adults in the age range 55-69 [109]. Respondents perceived that they had the skills to obtain, evaluate and use online health information however they were less assured in their ability to apply this information to a health choice. Moreover, information overload was also an issue, with respondents unsure of how to identify appropriate information from the volume of information available.

The research studies described used a range of survey instruments, interviews and Internet based performance tests to investigate the determinants and outcomes of health literacy and eHealth literacy. In common with these studies, this research utilises validated instruments to measure health and eHealth literacies, in addition to Internet based search tasks to scrutinise actual OHISB. However, the main novelty in the research carried out in this PhD involves the use of tracking software and an algorithm to monitor, parse and analyse the user's online interactions, search strategies and the types of information resources utilised. These behaviours are examined in order to determine associations between these and both health literacy and eHealth literacy.

3.3 Methodology

During the study each participant attended an online health information seeking experiment. Fig. 3.1 indicates the protocol followed during an experiment. In the first stage, the participant used a browser to search for information to answer six health questions. Following this they completed a health literacy instrument, the Newest Vital Sign [359], and an eHealth literacy instrument, the eHealth Literacy Scale [360]. In the

final stage they added demographic information including age, gender and highest qualification.

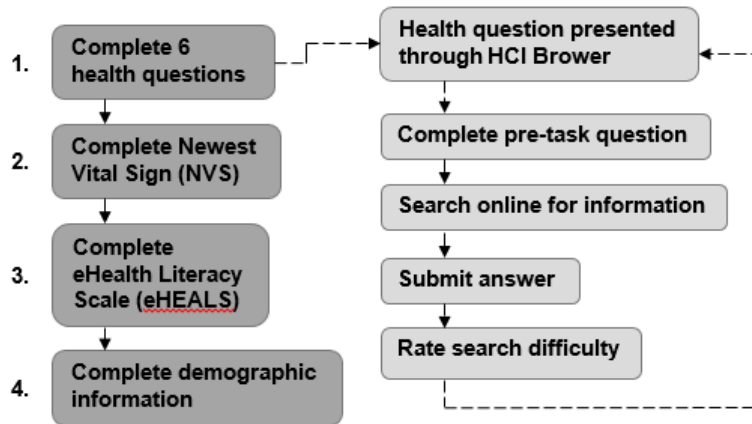


Fig. 3.1. Protocol for online health information seeking experiment.

3.3.1 Participants

A total of 60 subjects participated in the study, of whom 60% were male ($n = 36$) and 40% female ($n = 24$). Hence, a total of 360 online health information seeking tasks were recorded (6 tasks * 60 subjects). The age range of the group was 18 to 59, with a mean age of 27.43 years ($SD = 9.60$). The group included undergraduate and postgraduate university students and university staff, the majority of whom (61.7%, $n = 37$) had achieved an undergraduate or postgraduate degree, and the remaining participants (35%, $n = 21$) had a Regulated Qualifications Framework qualification at level 3 or level 4 [361]. Two participants did not enter any educational information. The background disciplines of the participants included Computer Science (63.3%, $n = 38$), Arts (10%, $n = 6$), Life and Health Sciences (6.7%, $n = 4$), Business (6.7%, $n = 4$), and Social Science (3.3%, $n = 2$). Six participants (10%) did not list a discipline.

3.3.2 Health Literacy

The Newest Vital Sign (NVS) [359] was used to measure a participant's level of health literacy. This is a health literacy screening instrument which provides a nutritional label, accompanied by six questions that measure literacy and numeracy skills. For this study the NVS-UK [362] was used. This is a validated version of the NVS in which the measurement scales and terminology used are consistent with UK nutrition labels. Each question is scored as correct or incorrect, resulting in a final sum score out of six. The

final score is used to classify a person's health literacy skills; a score of four or above indicates adequate health literacy, a score of two or three indicates intermediate health literacy, whilst a score of one or zero indicates low health literacy. A copy of the NVS instrument is available in Appendix A.

3.3.3 Ehealth Literacy

The eHealth Literacy Scale (eHEALS) [360] was used to quantify eHealth literacy. This instrument contains eight items through which a subject self-rates their ability to obtain, appraise and use electronic health information. The items are scored on a five-point Likert scale of one to five (ranging from strongly disagree to strongly agree) with a final sum score within the range of eight to 40. A higher end score indicates a higher level of eHealth literacy. Two supplementary items are also provided and can be used to evaluate a subject's general interest in using health resources on the Internet. These items were also completed by each participant. A copy of the eHEALS instrument is available in Appendix B.

3.3.4 Health Questions

During an online health information seeking experiment each participant was presented with six health questions and could search online for information to answer the questions. Before beginning a search the participant was asked to indicate whether they could answer the question without searching online, and could submit an answer directly if they wished. The health questions, which are listed in Table 3.1, focused on strategic areas such as diabetes, obesity, influenza, nutrition and analgesic medication. Each submitted answer was scored as correct or incorrect, with a final sum score out of six.

Table 3.1. Health questions presented to a participant during an online health information seeking experiment.

	Health Question	Context
1	You have a headache and buy a packet of Co-Codamol tablets from the chemist. Which are the two largest ingredients in Co-Codamol tablets?	Co-Codamol is one of the most frequently prescribed analgesic medications by General Practitioners in Northern Ireland [364].
2	Body mass index (BMI) is a measurement that indicates whether your weight is appropriate for your height. During a visit to your physician your BMI is determined to be 27. What category does your BMI fall into?	The prevalence of obesity in the UK is 29.8% [365]. The BMI index is commonly used to classify obesity and overweight in adults [366].
3	Diabetic Retinopathy is an eye complication associated with Diabetes. Can you name 3 common symptoms of this condition?	The prevalence of diabetes in the UK is 7.7% [365]. Diabetic retinopathy is the most common cause of blindness among people of working age in the UK [367].
4	The 5 A DAY message in the UK encourages people to eat five portions of fruit and vegetable every day. If you were to eat 5 portions of vegetables, of the recommended weight, in one day, how many grammes of vegetables would you eat in total?	It is a UK government recommendation that adults eat five portions of fruit and vegetables each day [368].
5	One widely publicised UK health message is that in most cases antibiotics should not be used to treat a common cold. Why is this advice given?	It is a widely promoted health message in the UK that antibiotics should not be used to treat a common cold [369].

6	Hypoglycemia (low blood glucose levels) is a complication of Diabetes. The unit of measurement for blood glucose level is Millimoles per litre (mmol/l). In terms of this measurement how is Hypoglycemia defined in the UK?	Self-monitoring for symptoms of Hypoglycemia is a recommended treatment for diabetes [370].
---	--	---

It can be difficult to accurately quantify OHISB therefore it was decided to integrate software into the web browser to objectively monitor online interactions, including search queries and Uniform Resource Locators (URL). The health questions were presented using the HCI Browser [363], which is a Mozilla Firefox extension that presented each question in turn and enabled the participant to implement their own search strategy, for example by using a search engine or entering a URL directly into the address bar. The HCI Browser also provided an interface through which an answer could be submitted, and collected timestamped browser log event data including the pages loaded, links clicked and the opening and closing of tabs. Fig. 3.2 (a) presents a screen capture of the HCI Browser, and (b) an example log event data file. A screen capture was also recorded as a participant worked through the six health questions. This enabled the collection of data that was not recorded by the HCI Browser, such as the rank order of results on a search engine results page (SERP). On completing a question the participant could rate, on a scale of one to five (very easy to very difficult), how difficult it had been to locate the information necessary to answer the question.

A new Java based parsing program was also developed to process the HCI Browser log files, and output these as a spreadsheet which could be utilised for further statistical analysis. The output included the search queries entered, the URL of visited web pages and the duration of each visit. Fig. 3.2 (c) presents an excerpt of a spreadsheet record produced by the HCI Parser. Pseudo code for the HCI Parser is available in Appendix C.

(a)

HCI Browser interface presenting health question

Task #2: Body mass index (BMI) is a measurement that indicates whether your weight is appropriate for your height. During a visit to your physician your BMI is determined to be 27. What category does your BMI fall into?

For this task, you may submit up to 1 answers. You have submitted 0 answers for this task.

Submit this answer

No answer - go back

Copy selected text to answer box

URL of answer

http://www.nhs.uk/livewell/loseweight

Answer text

Overweight

URL of answer source

http://www.nhs.uk/livewell/loseweight

Submitted answer

NHS choices

Your health, your choices

Health A-Z

Live Well

Care and support

Health news

Services near you

What's your BMI?

Share:

Enter a search term

(b)

Date	Time	Task type	Event	Online searching – URL of visited web page			
31-03-2015	16:14:05	S24	P24	T2	intask	Focus	http://www.nhs.uk/livewell/loseweight/
31-03-2015	16:14:06	S24	P24	T2	intask	Scroll	http://www.nhs.uk/livewell/loseweight/
31-03-2015	16:14:24	S24	P24	T2	intask	submittedAnswerURL	http://www.nhs.uk/livewell
31-03-2015	16:14:24	S24	P24	T2	intask	submittedAnswerText	Overweight http://www.nhs
31-03-2015	16:14:32	S24	P24	T2	posttask	onAccept event	
31-03-2015	16:14:32	S24	P24	T2	posttask	q1	2 null
		Participant	no	Health question	no	Post task search difficulty rating	Health question – submitted answer and source URL

(c)

Participant no	Health question no	Post task search difficulty rating	Duration of health question in seconds	Total answers submitted	Health question answer	Domain of answer source	URL of answer source
24	2	1	2	121	1	Overweight	www.nhs.uk http://www.nhs.uk/livewell/loseweight/pages/bodymassindex.aspx

Part	Que	PreTask	PostTask	Duration	TotalAnswers	Answer1	Ans1Domain	Ans1URL
24	2	1	2	121	1	Overweight	www.nhs.uk	http://www.nhs.uk/livewell/loseweight/pages/bodymassindex.aspx

SearchQueries	Query1	Query2	Ws1Domain	Ws1Duration	Ws1Type	Ws1Event	Ws1URL
2	BMI=27 AND category	BMI=27	www.google.co.uk	33	SearchResults	Scroll	http://www.google.co.uk/search?q=BMI%3D27&ie=utf-8

Total search queries entered during health question

Search queries

Domain of first web page loaded during health question

Web page visit length in seconds

Web page type (SERP)

Scroll event occurred

Web page URL containing query strings

Fig. 3.2. HCI Browser and HCI Parser components, (a) screen capture of HCI Browser interface for submitting a health question answer, (b) log file containing records for web page visit, health question answer and post-task question choice, and (c) excerpt of spreadsheet record for health question including duration, answer, search queries and first web page visited during searching.

3.3.5 Statistical Analysis

Bivariate analysis of the data was conducted using SPSS version 23. Descriptive statistics were used to calculate the mean and standard deviation of the NVS, eHEALS and health question scores, and Pearson's product-moment correlation was used to evaluate the correlations between these scores. Multiple regression analysis was also used to determine whether demographic variables were significant predictors of NVS and eHEALS scores. Chi-Square analysis was also used to investigate the proportions between various characteristics of OHISB.

3.4 Results

3.4.1 Health Literacy, Ehealth Literacy and Health Questions

The NVS scores indicated that 75% ($n = 45$) of the participants had adequate health literacy skills, 16.7% ($n = 10$) had intermediate health literacy skills, and 8.3% ($n = 5$) had low health literacy skills. The mean NVS score was 4.53 (SD 1.73) out of six. Analysis of the internal consistency of eHEALS found high reliability, $\alpha = .859$. The eHEALS scores ranged from 17 to 40, with a mean score of 29.9 (SD 5.68). Fig. 3.3 indicates the frequencies of the NVS and eHEALS scores. The correlation between the NVS and eHEALS scores was not statistically significant ($r = .144$, $p = .272$). The relationship between the two scores is illustrated in Fig. 3.4. The results from the eHEALS supplementary items suggested that most of the participants had a positive perception of Internet based health resources; 76.7%, ($n = 46$) indicated that the Internet was useful in helping them make decisions about health, and 86.6% ($n = 52$) indicated that it was important to them to be able to access health resources on the Internet. Moreover, many of the participants appeared confident in their ability to use the Internet as a source of health information; 83.3% ($n = 50$) agreed that they knew how to use the Internet to answer health questions, and 66.7% ($n = 40$) agreed that they had the skills required to evaluate health resources found on the Internet.

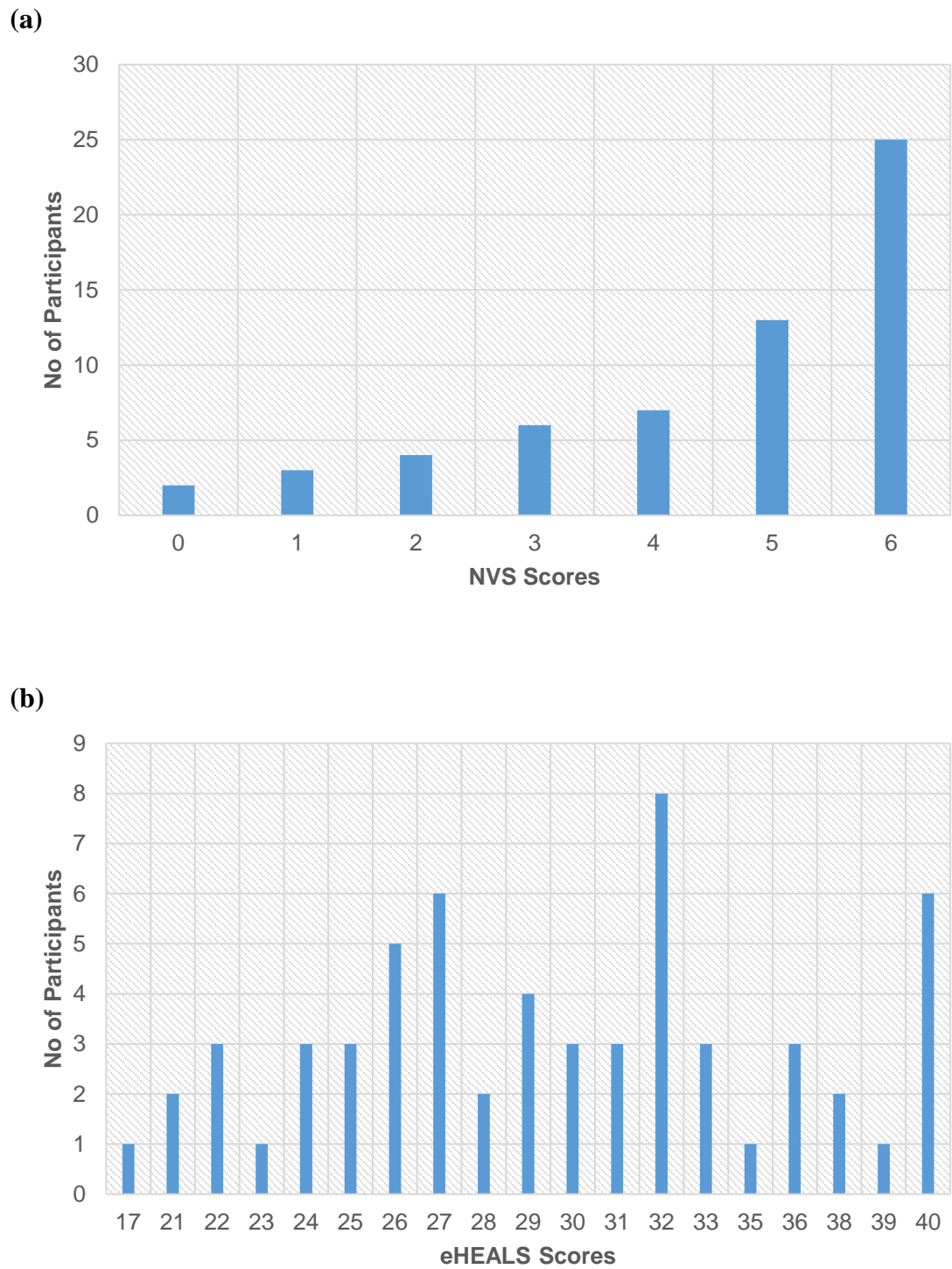


Fig. 3.3. Bar charts indicating frequencies of (a) NVS scores and, (b) eHEALS scores.

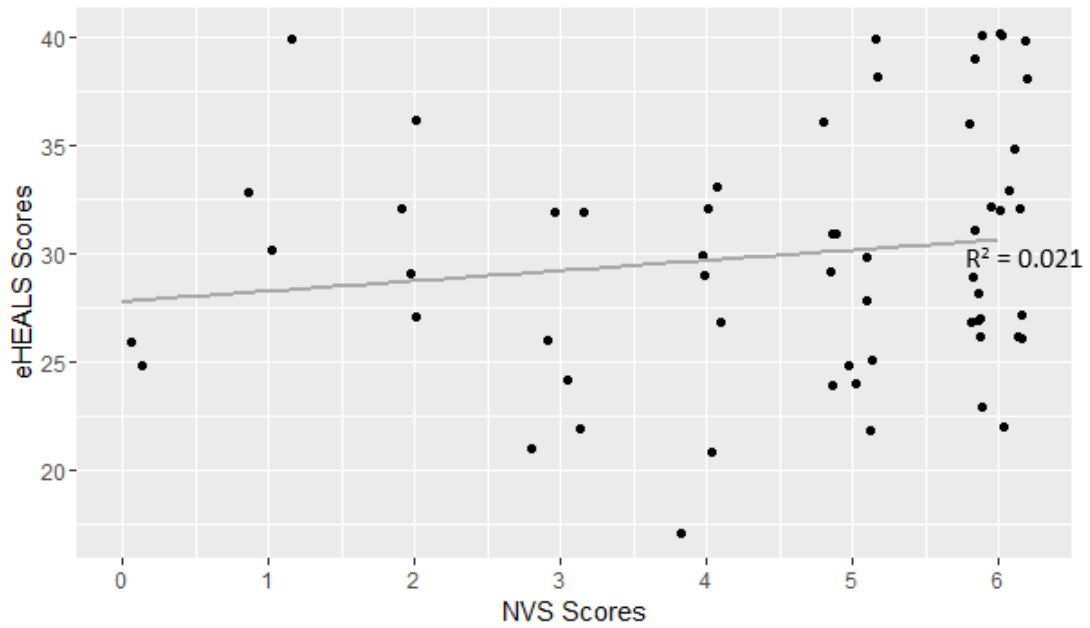


Fig. 3.4. Scatterplot with linear regression line illustrating correlation between NVS scores and eHEALS scores.

Several studies have identified age and educational attainment as determinants of an individual's level of health literacy and eHealth literacy, therefore two regression models were developed in order to investigate whether these characteristics had any predictive ability on the health literacy and eHealth literacy scores of the study population. Table 3.2 shows the results of multivariate regression conducted using NVS scores as the dependent variable, and qualification level and age as explanatory independent variables. The regression model represented a statistically significant proportion of the variance ($R^2 = .180$, $F = 6.265$, $p = .003$), however only qualification level was a significant predictor of NVS score. Multivariate regression analysis was also carried out for the eHEALS scores however the variance explained by the model was not statistically significant ($R^2 = .060$, $F = 1.824$, $p = .171$).

Table 3.2. Results of multiple regression for NVS scores.

	B	SE B	β	p
Constant	2.824	.711		
Qualification Level	.367	.108	.437	.001
Age	-.007	.023	-.039	.765

$R^2 = .180$

Fig. 3.5 indicates the health question score frequencies. All participants were able to answer at least one of the health questions correctly. The majority of subjects (78.3%, $n = 47$) answered four or more of the health questions correctly, 15% ($n = 9$) gained a full score of six and 3.3% ($n = 2$) answered only one question correctly. The mean score was 4.22 ($SD = 1.26$). Multivariate regression analysis was conducted to evaluate whether age and educational attainment were predictors of health question score. Table 3.3 shows the results of this analysis. The regression model represented a statistically significant proportion of the variance ($R^2 = .194$, $F = 6.877$, $p = .002$), however again only qualification level was a significant predictor of question score.

Table 3.3. Results of multiple regression for health question scores.

	B	SE B	β	p
Constant	2.737	.514		
Qualification Level	.262	.078	.428	.001
Age	.004	.017	.034	.793

$R^2 = .194$

Pearson's product-moment correlation was used to evaluate whether any associations existed between the question score achieved and a participant's level of health literacy, eHealth literacy or highest qualification. Health question scores had a statistically significant weak positive correlation with NVS score ($r = .34$, $p < .01$) and qualification level ($r = .41$, $p < .01$). However, there was no significant correlation between question score and the eHEALS score. These correlations have also been presented in the form of scatter plots in Fig. 3.6.

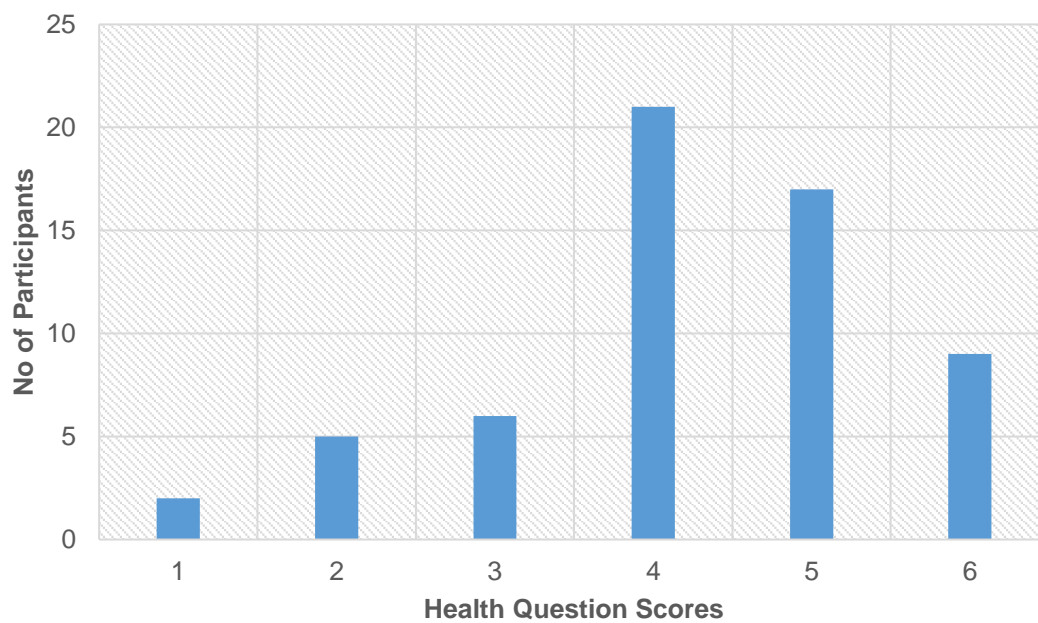
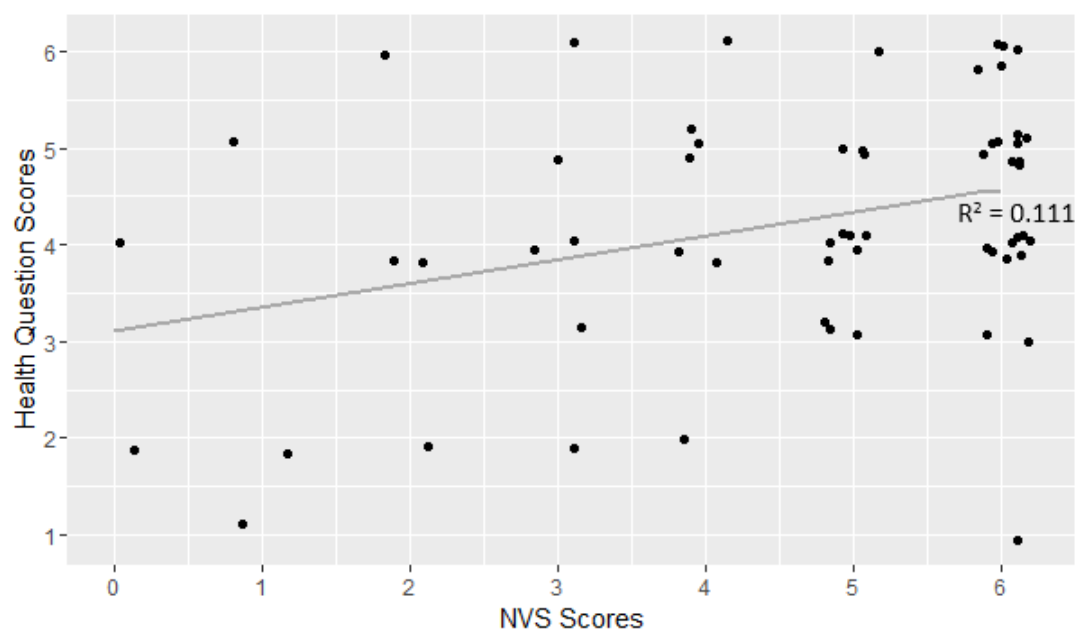
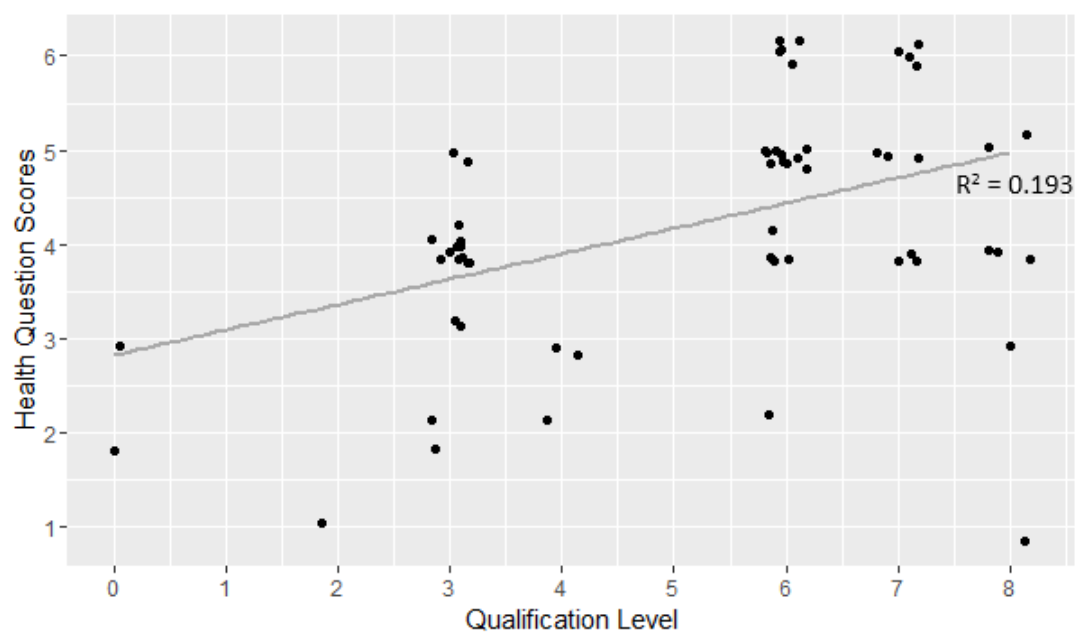


Fig. 3.5. Frequencies of health question scores for study population.

(a)



(b)



(c)

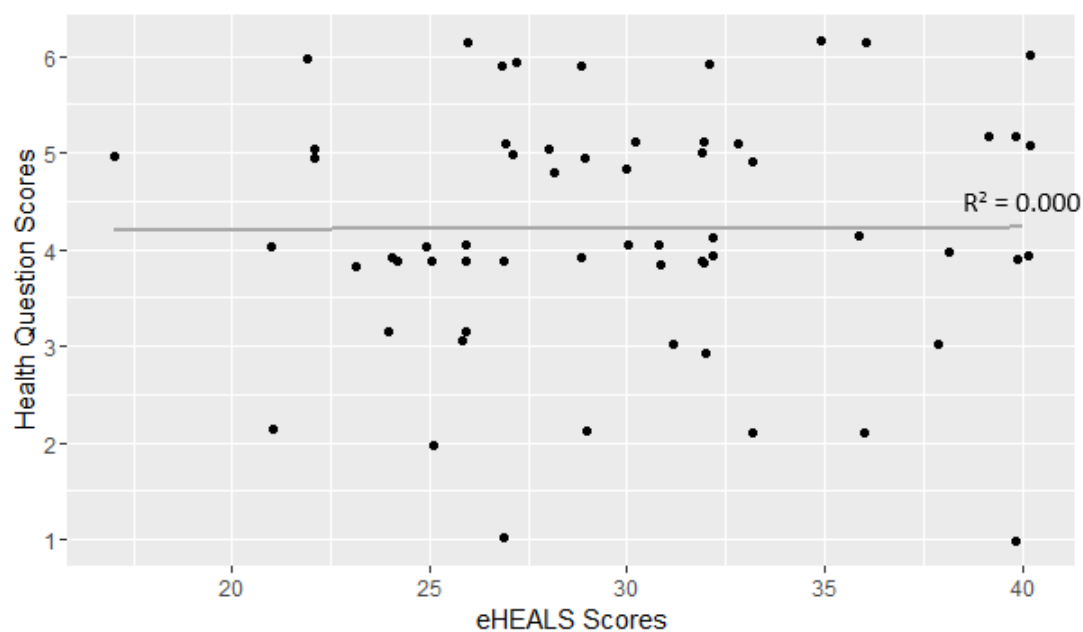


Fig. 3.6. Scatterplots with linear regression line illustrating correlations between (a) question score and NVS score, (b) question score and qualification level, and (c) question score and eHEALS score.

3.4.2 Health Information Seeking Behaviour

Fig. 3.7 indicates the total submitted and correct answers for each question, and Fig. 3.8 indicates the average completion times for correct and incorrect answers. Table 3.4 illustrates paired *t*-test analysis of the completion times for all the health questions. As can be perceived, there was variation in the completion rate and the time taken to complete each health question. The correlations between completion times and question scores were not statistically significant for any of the health questions. Furthermore, as denoted in Fig. 3.8, there was no significant difference between the completion times for correct and incorrect answers for any of the questions. Question 4 had the lowest completion rate ($n = 50$) and the lowest proportion of correct answers ($n = 29$). Moreover, as indicated in Table 3.4 the mean completion time for question 4 was significantly longer than all other questions with the exception of question 6. Question 3 had the highest rate of correct answers ($n = 54$) and a Chi-Square test of independence indicated that, with statistical significance, the participants performed best in this question, $X^2(5) = 42.515, p < .001$.

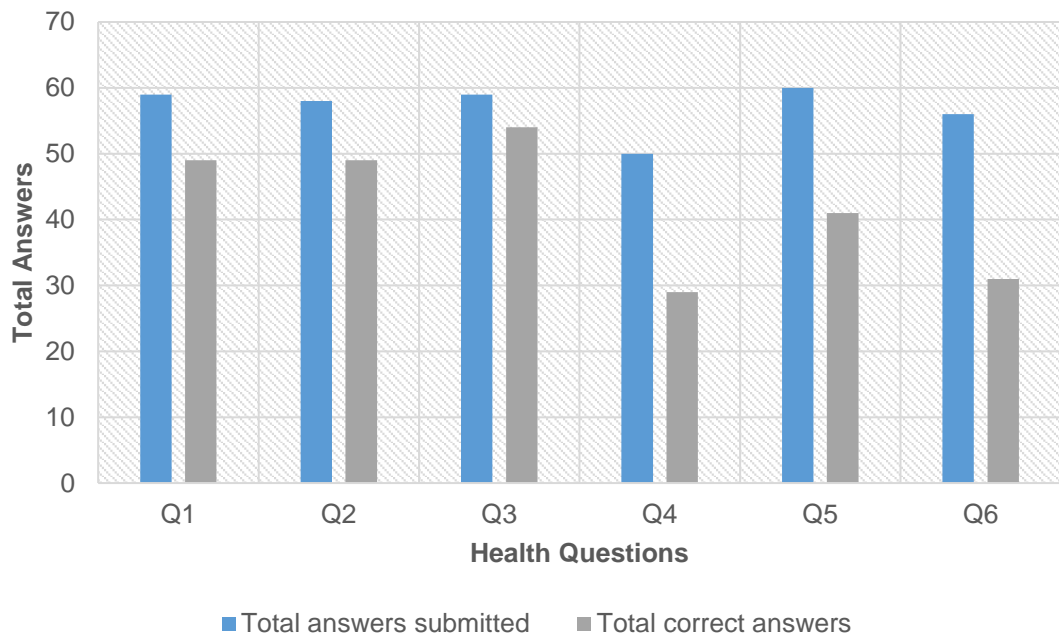


Fig. 3.7. Total submitted and correct answers for health questions.

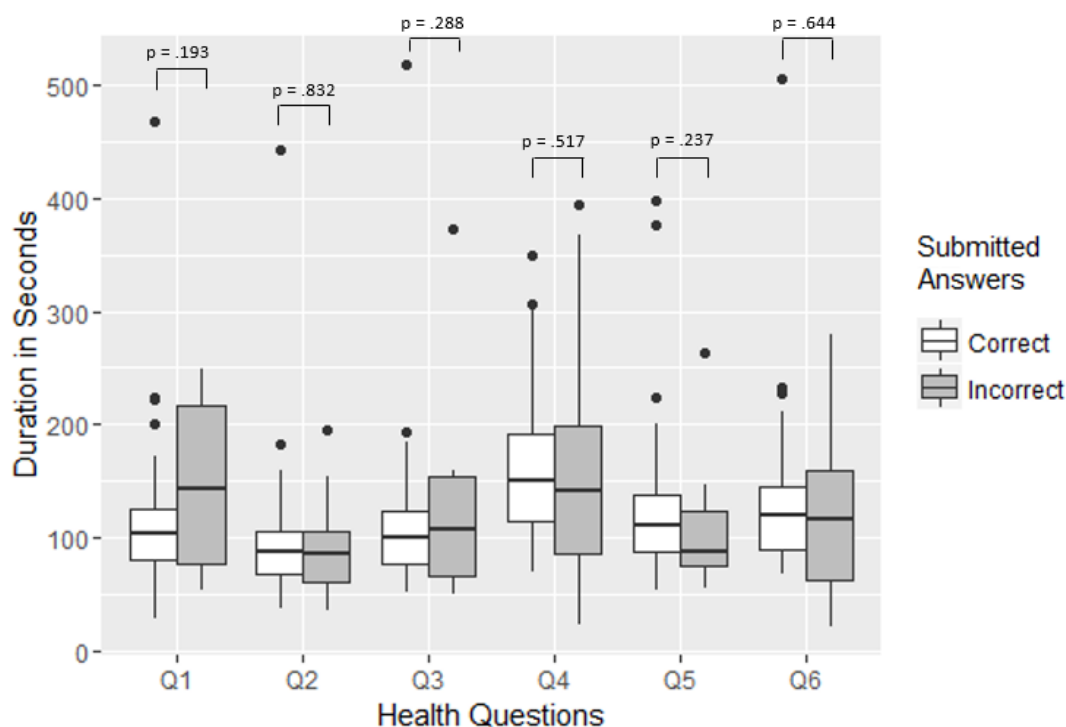


Fig. 3.8. Boxplots of completion times for correct and incorrect answers, including p -values from t -test comparison.

As the health questions related to popular health topics and had been designed to be of equal difficulty it was unclear why there was such variation in the submission and accuracy rates. Therefore, it was decided to explore the health information seeking behaviours exhibited throughout the experiment in order to determine whether behaviours differed between questions. The specific aspects of OHISB which would be analysed included, (1) time spent on visits to SERPs and information web pages, (2) the types of web resources visited during searching activities, and utilised as answer sources, (3) the search queries entered, and (4) accessing of search results.

3.4.2.1 Time Spent on SERPs and Information Pages

The most prominent method for seeking information throughout the experiment was through a search engine. As there was such a reliance on SERPs to find sources of health information t -test analysis was conducted to determine if there were any distinctions between the time spent on SERPs during each of the six health questions. The results of this analysis are illustrated in Table 3.5.

Table 3.4. Results of *t*-test comparison of the completion times for the health questions, indicating *t* statistic and *p*-value.

	Q1	Q2	Q3	Q4	Q5	Q6	Mean	SD
Q1		3.63**	.85	-2.75**	.13	-1.06	120.75	68.21
Q2	-3.63**		-2.40*	-6.49****	-3.12**	-3.80****	94.70	56.88
Q3	-.85	2.40*		-3.72****	-.75	-1.83	113.10	72.55
Q4	2.75**	6.49****	3.72****		3.17**	1.97	157.55	79.34
Q5	-.13	3.12**	.75	-3.17**		-1.16	119.47	65.26
Q6	1.06	3.80****	1.83	-1.97	1.16		131.02	78.71

p*< .05, *p*< .01, ****p*< .001 (two tailed).

Table 3.5. Results of *t*-test comparison of SERPs visit duration during health questions, indicating *t* statistic and *p*-value.

	Q1	Q2	Q3	Q4	Q5	Q6	Mean	SD
Q1		2.95**	4.62***	-.52	1.95	.57	23.23	21.65
Q2	-2.95**		1.50	-3.09**	-.57	-2.46*	14.90	12.94
Q3	-4.62***	-1.50		-4.04****	-1.67	-3.80****	11.23	16.00
Q4	.52	3.09**	4.04***		2.32*	1.01	25.32	25.14
Q5	-1.95	.57	1.67	-2.32*		-1.52	16.42	22.41
Q6	-.57	2.46*	3.80****	-1.01	1.52		21.47	19.25

p*< .05, *p*< .01, ****p*< .001 (two tailed).

Question 4 had the largest total of time spent on visits to SERPs (mean time per subject = 25.32 seconds, SD = 25.14), and the mean visit duration was significantly longer than the mean time for questions 2, 3 and 5. Question 3 had the smallest total of time spent on SERPs (mean time per subject = 11.23 seconds, SD = 16.00), and this was significantly smaller than the mean time for questions 1, 4 and 6. However on further investigation it became apparent that the reduced time on SERPs during question 3 could be attributed to the OHISB of the participants. The majority of participants (90.0%, n = 54) only visited one SERP and found a link to the source that they used to answer the question, whilst five participants (8.3%) visited only two SERPs before locating an answer source. In comparison, the total number of SERPs visited during question 4 was much larger. Again most of the participants (81.7%, n = 49) found an answer source from the first or second SERP visited, however the remaining participants (15.0%, n = 9) carried out at least three distinct searches and visited at least three distinct SERPs during their searching activities; 8.3% (n = 5) visited three distinct SERPs, 5.0% (n = 3) visited four distinct SERPs and one participant (1.7%) visited five distinct SERPs. The higher rates of distinct searches during question 4 suggests that participants may have found it more challenging to find the information they sought than during question 3. The mean SERP visit duration for question 4 was more than double the mean duration for question 3, implying that on average participants spent substantially longer reading through the search results during question 4. Again this suggests that participants may have found it more problematic to identify suitable sources of information from the search results during question 4. The scrolling behaviour exhibited during question 4 also suggests that many participants read through a larger number of the search results than they did in question 3. During question 3 the majority of participants (73.3%, n = 44) either used information from the first SERP visited to answer the question, or chose an answer source from the first four results presented and did not scroll down the SERP to look at any further results. In total only 13 participants (21.6%) scrolled beyond the first four SERP results. In comparison 32 participants (53.3%) scrolled beyond the first four results on at least one of the SERPs that they visited during question 4. This would suggest that many participants found it more demanding to locate suitable information resources during question 4 and scrutinised the search results more extensively than they did in question 3.

Table 3.6. Results of *t*-test comparison of information page visit duration during health questions, indicating *t* statistic and *p*-value.

	Q1	Q2	Q3	Q4	Q5	Q6	Mean	SD
Q1		2.03*	-1.15	-2.07*	-.34	-.08	44.53	50.11
Q2	-2.03*		-3.07**	-3.76***	-2.15*	-1.88	33.55	37.68
Q3	1.15	3.07**		-1.36	.75	1.11	53.32	56.21
Q4	2.07*	3.76***	1.36		1.99	2.05*	66.40	63.88
Q5	.34	2.15*	-.75	-1.99		.28	47.25	49.77
Q6	.08	1.88	-1.11	-2.05*	-.28		45.02	56.09

* $p < .05$, ** $p < .01$, *** $p < .001$ (two tailed).

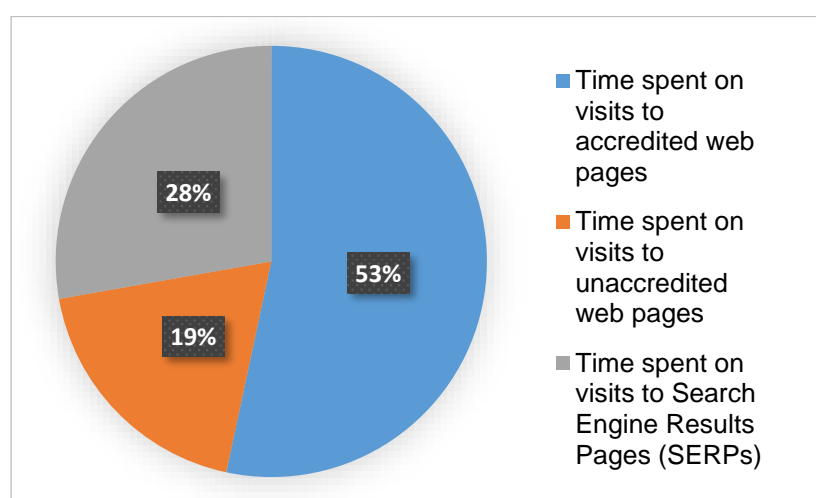
Throughout the experiment the participants visited different sources of information including government, academic and commercial websites, accredited sources of health information and unaccredited sources of health information such as blogs. Question 4 had the largest sum total of time spent on information web page visits (mean time per subject = 66.40 seconds, SD = 63.88) and question 2 had the smallest (mean time per subject = 33.55 seconds, SD = 37.68). Table 3.6 provides the results of *t*-test comparison of the mean time spent on information pages during the experiment. Again no clear patterns are apparent. Although there is a significant time difference between question 4 and questions 1, 2 and 6 there is no significant difference with questions 3 and 5. The mean time for question 2 is significantly smaller than the mean time for all questions except for question 6. An investigation of the OHISB of the participants during question 2 clarified the reasons why the mean visit duration was low. In total 26 participants (43.3%) visited only one information web page, and used this source to answer the question. In addition, 18.3% ($n = 11$) used a SERP to answer the question and did not visit any information pages at all. It was surprising to note that question 2 had the second highest rate of visits to information pages ($n = 66$) whilst question 4 had the highest rate ($n = 71$). However, the mean information page visit duration during question 2 was approximately half of the mean duration for question 4. This implies that on average participants spent twice as long reading an information page during question 4 than they did in question 2, again suggesting that participants found it problematic to locate the information required for question 4.

With regard to the overall distribution of the time online during the study, the majority of time (72%) was spent on information pages whilst 28% of the time was spent on SERPs. Fig. 3.9 (a) indicates the proportions of time that were spent on visits to accredited and unaccredited information web pages, and SERPs. In total there were 814 distinct web page visits. As indicated in Fig. 3.9 (b) SERPs were the most frequently visited artefact, however Fig. 3.9 (a) specifies that, in comparison, visits to accredited information web pages constituted a larger proportion of the time spent online (53%). At 47.1 seconds, the average visit duration for an accredited web page was more than double the average visit duration for a SERP (15.5 seconds per visit). Moreover, the average SERP visit duration was also less than the average duration for an unaccredited web page (41.7 seconds). Thereby although SERPs had a higher rate of visits (432 distinct SERPs visited) than accredited web pages (273 distinct web page visits) and

unaccredited web pages (109 distinct web page visits), SERP visits tended to be shorter, implying that on average the participants spent much longer scrutinising information than they did search results.

The prominent method used to find information was through SERPs, however there were no statistically significant correlations between NVS and eHEALS scores and the average amount of time that a participant spent on visits to SERPs, or accredited or unaccredited information pages.

(a)



(b)

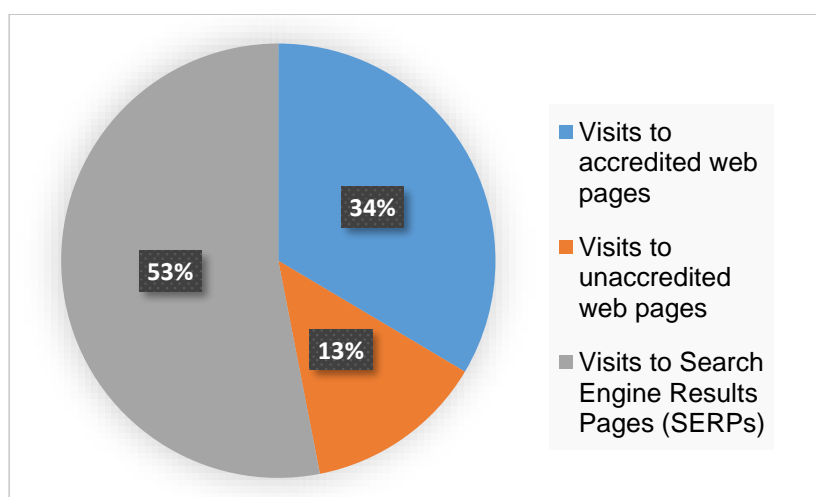


Fig. 3.9. Pie charts indicating (a) percentage of total time spent on visits to different categories of web resource, and (b) percentage of total distinct visits to different categories of web resource.

Table 3.7. Ten most frequently visited domains throughout the study, with brackets indicating rank order of total visits and use as an answer source.

Domain	Accredited source	Total number of visits to domain	Total submitted answers which used domain as source
google.co.uk		395 (1)	59 (2)
nhs.uk	✓	151 (2)	109 (1)
diabetes.co.uk		41 (3)	30 (3)
google.com		37 (4)	6 (9)
en.wikipedia.org		26 (5)	17 (5)
webmd.com	✓	25 (6)	20 (4)
medicines.org.uk	✓	19 (7)	14 (6)
patient.co.uk	✓	15 (8)	12 (7)
webmd.boots.com	✓	15 (8)	8 (8)
apps.who.int	✓	8 (9)	4 (10)

In total 53 distinct domains were visited throughout the study, however the participants tended to revisit particular domains more frequently than others. Table 3.7 indicates the 10 most frequently visited domains. Visits to these domains constituted 89.9% of the total page visits, and moreover, these domains comprised the answer sources for 81.6% of the submitted answers. As expected a search engine domain, google.com, was the most frequently visited, and had more than double the visits of the second most visited domain nhs.uk. However, in contrast, nhs.uk was the most frequently used source for health question answers whilst google.com was the second most commonly used source.

3.4.2.2 Health Question Answers

In total 342 answers were submitted, of which 74% (n = 253) were correct. Over half of the total submitted answers (55.8%, n = 191) used accredited sources of information and 80.6% (n = 154) of these answers were correct. A much smaller proportion of the submitted answers (20.2%, n = 69) used unaccredited websites and 71% (n = 49) of these answers were correct. There was a notable use of SERPs as an information source;

18.7% ($n = 64$) of the submitted answers were gathered from SERPs, and over half of these answers (67.2%, $n = 43$) were correct. The results of a Chi-Square test of independence indicated a significant association between the information source (accredited, unaccredited or SERP) and whether the submitted answer was correct, $X^2(2) = 7.721, p < .05$. Only two participants (3.3%) used accredited sources to answer all six questions. There was no distinct pattern as to the usage of SERPs to gather information. Over half of the participants (56.7%, $n = 34$) used information from a SERP to answer at least one question, while 3.3% ($n = 2$) used SERPs to answer five out of six health questions. Investigations to determine whether there were any correlations between NVS score and the types of resources used determined a statistically significant weak positive correlation between NVS score and the use of accredited sources to answer the health questions ($r = .27, p < .05$). A similar relationship was also found between qualification level and the use of accredited answer sources ($r = .26, p < .05$). However, no significant correlation was found between eHEALS scores and the use of accredited sources. A statistically significant weak negative correlation was identified between NVS score and the use of unaccredited sources to answer the health questions ($r = .29, p < .05$). No significant relationships were identified between the number of answers from unaccredited information sources or SERPs, and eHEALS scores, age or qualification level.

A notable trait amongst the study population was that, in general, the participants did not compare information between different sources. For 57% ($n = 195$) of the submitted answers the participant used information from the first information page that they visited and did not visit any other information resources. Furthermore, for 16.4% ($n = 56$) of the submitted answers the participant gathered information from a SERP and did not visit any additional SERPs or any information pages at all.

3.4.2.3 Search Queries

Search query formulation is an increasing focus of OHISB research [136]. The search queries constructed by an information seeker can reflect how they understand and express their health needs. Moreover, search query reformulations can reveal how they reconceptualise their health information needs in response to information gathered or challenges faced during searching activities. As the reliance on search engines was so pronounced during the study it was decided to examine the search queries entered. The

analysis focused on search query length and vocabulary as this would reflect how the participants conceptualised and articulated their information needs during the health questions.

During the study the participants entered 426 search queries which comprised 2594 distinct terms. There was disparity in the length of the search queries entered, and question 2 had the shortest average length (3.1 terms), and question 4 the longest average length (9.9 terms). The average length for the combined search queries was 6.1 terms. Examination of the search queries revealed that 17.4% ($n = 74$) contained one or more spelling mistakes. However, there was a high level of repetition amongst the misspelt terms, with one prominent example being the term *co-codamol* which was misspelt in 27 distinct search queries, whilst the term *symptoms* was misspelt in 12 distinct queries. In total the search queries contained 99 misspelt terms, and 87 (87.9%) of these were key terms that had been included in the health questions. There was a notable use of stop words (e.g. and, are) and interrogative words (e.g. why, what) in the search queries, however only one search query contained the Boolean operator AND. A prominent feature of the search queries was the frequent use of key terms drawn from the health questions. Over half of the total terms used (61.9%) had been obtained directly from the health questions, or were misspellings, alternative spellings or synonyms of these terms. Table 3.8 indicates the 20 most frequently occurring key terms in the search queries, all of which were drawn from the health questions.

There were 88 reformulations of search queries, of which 35.2% ($n = 31$) were spelling reformulations. There was no clear pattern regarding the addition, removal or substitution of terms, however several queries were narrowed by adding specific terms of measurement (e.g. millimoles, mmol/l, grammes), or location (UK) which were taken directly from the corresponding health question.

Table 3.8. Twenty most frequently occurring search query terms in rank order, including synonyms and misspellings of term.

Term	Total occurrences of term	Total occurrences of synonyms for term	Total occurrences of misspellings of term
day	76		
hypoglycemia	71		5
5	67		
vegetables	67	1	8
ingredients	63	1	3
retinopathy	61		1
diabetic	57		1
bmi	54		1
antibiotics	50	1	2
uk	50	2	
cold	49	4	1
common	48		1
symptoms	48		12
co-codamol	43		27
weight	41		2
grammes	31	36	1
portions	30	11	
27	29		
treat	29	1	
measurement	22	4	

3.4.2.4 Accessing Search Results

During the study 382 information web pages were visited, of which 72.5% (n = 277) were visited directly from a SERP, whilst the remainder (27.5%, n = 105) were accessed through hyperlinks within a website. All of the SERP links followed were chosen from the first page of search results and none of the participants visited any of the following SERPs. Of the links that were chosen from a SERP, the majority (96.4%, n = 267) were

chosen from the top five results, however the participants did not always select links in the order as presented. For example, the most frequently chosen domain from SERP results was nhs.uk ($n = 136$). However, on 63 distinct occasions a web page from the domain nhs.uk was the first link selected by the participant even though it was not the first result in the results list.

3.5 Discussion

The aim of this study was to discover whether there are any relationships between how an individual obtains and utilises online health information, and their level of health literacy and eHealth literacy. There was a high level of educational attainment amongst the participants and most of the participants had adequate health literacy scores. Although there was a wider distribution in the eHealth literacy scores over half of the study population perceived their skills as being in the upper quadrant. Attitudes towards Internet based health resources were largely positive as most of the participants perceived these as useful and convenient. Participants were also confident in their ability to use the Internet to gather health information. Education was found to be a significant predictor of health literacy, and this finding concurs with the results of similar research in the domain of health literacy [154, 156]. However, surprisingly neither education nor age were significant co-variants of the eHEALS scores, and no significant relationship was identified between the NVS and eHEALS scores. As health literacy is one of the core literacies of eHealth literacy it was unexpected that no correlation was found between the two scores. Moreover, whilst a weak positive correlation was found between health question scores and NVS scores, no relationship was found between the question scores and eHEALS scores. Again this was surprising as higher eHEALS results have been associated with enhanced online searching skills, and success at finding and using health information [351]. The eHEALS scores suggested that most of the participants had adequate skills to find the information necessary to answer all the health questions. However, the lack of correlation with the question scores suggests that the eHEALS scores may not have accurately reflected the actual online health information seeking abilities of all the participants. This finding corresponds with concerns that have been expressed regarding the validity of eHEALS as an instrument to quantify actual online health information seeking skills [154, 371]. Moreover it has also been suggested that as eHEALS is a measure of self-efficacy, the score may reflect an individual's overestimation of their skills rather than their actual

abilities [154, 371, 372]. However, it was not possible to quantify whether this was the case with the study population.

Although there was variation in the health question scores, all the participants were able to use online health information to correctly answer at least one health question. The low submission rate and manner of health information seeking suggested that question 4 was the most challenging for the participants. However, it was unclear why this question appeared so problematic. The majority of participants whom did not submit an answer or submitted an incorrect answer had adequate health literacy and eHealth literacy skills, thereby suggesting that they would have sufficient skills to locate and use the information required to answer this question. For a small proportion of participants, it appeared that appropriate information was gathered but an error was introduced whilst calculating the total number of grammes.

Though there was disparity in the participants' health question performance there were notable similarities in their OHISB. It had been emphasised to the participants that they were free to choose their own search strategy however the most frequent method for discovering health information was through search engine results. This behaviour is comparable to the searching activities of other online health information seekers [23, 110]. Throughout the experiment, participants prominently utilised SERPs to locate sources of information. In order to quantify the use of SERPs, the average SERP visit duration by a participant was measured. However, no significant correlation between SERPs average visit duration and health literacy or eHealth literacy was found. Employing a search engine based searching strategy did not appear to adversely affect information gathering as the majority of participants were able to locate sources to help them answer all six questions.

Search query formulation was characterised by a significant usage of key terms from the health questions. Moreover, one surprising feature of the search queries was the occurrence of misspelt terms, most of which had been available in the health questions. The incident of misspelt terms was surprising amongst such a highly qualified group of health information seekers. However, it is possible that the level of misspelling and reliance on health question vocabulary may have been associated with the participants' lack of knowledge on the related area. All of the participants that completed an online

searching task had indicated in the pre-task question that they did not have sufficient knowledge to answer the corresponding health question and therefore may have relied on the health questions to derive appropriate search query terms.

Despite having adequate health literacy and eHealth literacy skills most individuals utilised both accredited and uncertified health information. This included a significant proportion of participants who utilised information from SERPs snippets to answer the questions. The use of unaccredited health information sources and SERPs was surprising as 68.3% ($n = 41$) of participants had agreed that they could distinguish between high and low quality online health resources. However, although several participants used SERPs to answer more than one question, no significant predictors were found for this behaviour. Despite the recurrent use of uncertified health information, it must also be noted that the participants did tend to utilise a greater volume of accredited resources during searching activities and as answer sources. For example, the sum of distinct visits to accredited information web pages was more than double the total for unaccredited web pages. One notable behaviour was a tendency to favour the domain *nhs.uk*. This domain was both the most frequently visited, and the most often used as an answer source. Moreover, in a number of cases the participants purposely searched for NHS based information, as seven of the search queries contained the term *NHS*. No other information resource was directly sought in this way. A web page from the *nhs.uk* domain was chosen as the first link followed from a SERP more often than any other domain, and in many cases this link was chosen even though it was not the first result displayed on the SERP.

Another behaviour displayed by most of the participants was a reluctance to compare information between different sources. Almost all participants stopped searching at the page from which they had answered the question. For a large proportion of the questions, the answer source was the first page visited, thereby these participants made no effort to compare or validate this information with any other sources.

Many of the participants were highly confident in their online searching skills, however question scores and searching behaviours did not appear to reflect the eHealth literacy competencies for some participants. In order to further analyse and explicate some of the inconsistencies in the findings we should also consider how contextual factors may

have influenced the health information seeking behaviours of some of the participants. Online health information is often sought when health consumers perceive a risk or uncertainty regarding a personal or familial health situation [88, 373]. Health uncertainty often effects an emotional response such as anxiety or distress, and thereby the consumer may place greater value on the credibility and reliability of health information and adapt their search behaviours accordingly. Studies have suggested that when seeking information for salient health concerns online information seekers may undertake a more exploratory searching approach, performing multiple searches and examining various sources to compare and verify information, and placing greater significance on the reliability of the information source [25, 88, 120, 374, 375]. However, in contrast, the study participants did not have emotive motivations for searching for health information, and this may have influenced their engagement with the search tasks. As the topics under investigation were not personally relevant, some of the participants may have been less inclined to validate the information found, verify the reliability of the information source or persist in a challenging search situation. As a result, for some participants, the searching behaviours exhibited and question score achieved may not have accurately depicted their actual skills to locate and use online health information. This may have been reflected in some of the findings including the lack of correlation between the eHEALS scores and question scores.

The technical skills exhibited by the subjects suggests that this was a very capable group of online health information seekers. However, there are several possible confounding factors which must also be taken into account when considering the results of the study. The average age (approximately 27 years) and the high level of academic achievement among the subjects denotes that this was a predominantly younger and well-educated study population. Subsequently, it is likely that these attributes could have influenced the behaviour of the participants, and therefore the results, in some way. Although age was not identified as a significant predictor of health literacy, eHealth literacy or question score in this study, it is often found that younger and more educated population groups are characterised as having advanced health literacy [34-36, 376, 377]. Thereby, it is probable that being highly educated, most of the participants would have had advanced skills to understand, critique and use health information, a noted advantage in the experiments. The study population also exhibited adeptness in using online searching applications to locate information sources. Younger generational groups are

frequently identified as regular users of Internet based applications for a variety of activities including information search, social media usage and entertainment [378-381]. Subsequently, it would be likely that individuals of this age group would have experience of using online searching applications. Moreover, as over half of the study population had a Computer Science background, it would be expected that these subjects would have had advanced expertise in using online search applications, constructing search queries, and scrutinising search results. Another important factor which may have impacted the motivation of the participants was their information need. It is likely that as the information seeking was taking place as part of an experiment and not for a health concern, for some participants there was little saliency or interest in the information found. Thereby, as previously indicated, the participants may have paid less attention to the credibility of the information sources, and concentrated instead on quickly gathering the required information to complete the experiment. Conversely, as health information seeking is a common online activity, some of the subjects may have been especially interested in this topic which encouraged their motivation to perform well in the experiment. However, while it is likely that these factors may have influenced the behaviour of the participants during the study, the extent to which these factors did influence the results of the study remains uncertain.

The completion rate of the health questions suggests that the participants were highly engaged during the experiment. However, it is also important to consider whether the experimental conditions could have influenced the motivation of the subjects in any way. Although the participants were assured that they were not under direct observation by the investigator, it is possible that the presence of the investigator may have affected the behaviour of some participants. The *Hawthorne effect* relates to a situation whereby an individual, being aware that they are being observed, alters a pertinent aspect of their behaviour [382, 383]. The Hawthorne effect can occur during research experiments despite efforts being made by an observer to avoid influencing or affecting behaviour [382]. In many cases the behavioural changes are effected to match the perceived needs or expectations of the observer [382, 383]. The attendance of the investigator during these experiments may have affected the focus and motivation of the subjects in several ways. For example, being mindful of the investigator, some participants may have felt obliged to complete the whole experiment. Alternatively, as some of the participants were colleagues of the investigator, these subjects may have been especially keen to

succeed in the health searching tasks and to complete the experiment. However, whilst acknowledging these possible influences, it is difficult to gauge the extent to which the experimental conditions did effect the engagement and motivation of the participants.

3.5.1 Study Limitations

There were a number of limitations in this study. Firstly, the participants were drawn from a convenience sample of university students and staff and thus it was likely that the level of educational attainment would be high, and consequently health literacy and eHealth literacy levels would also be high for most of the study population. However, the findings show that there was diversity in the range of scores achieved in NVS and eHEALS. Moreover, although the study population perceived themselves to be advanced users of the Internet for health information seeking purposes, the study highlighted some limitations in the effectiveness of their searching behaviours. A second limitation was that the study was performed under laboratory conditions and it is likely that the participants completed the search tasks in less time than they would take if they were actually seeking information for a genuine health concern. However, presenting the participants with questions related to popular public health topics enabled the analysis of a range of OHISB that were typical of daily online health information seeking.

3.6 Conclusion

The research question queried whether any relationships exist between an individual's level of health literacy and eHealth literacy and their behaviours when seeking and using online health information. In order to answer this question, a research study was completed in which participants searched online for information to answer health questions. Subsequent analysis focused on examining the relationships between aspects of the participants' OHISB and their level of health literacy and eHealth literacy. Bivariate analysis identified statistically significant positive correlations between NVS scores and both health question scores and the use of accredited sources to answer the health questions. This would suggest that relationships did exist between health literacy and OHISB, and that an increase in health literacy correlated with increased skills to answer the health questions, and an increased likelihood of using accredited sources to answer the health questions. However, similar tests found no statistically significant associations between the eHEALS scores and any of the behaviours quantified, thereby

suggesting that no relationships exist between eHealth literacy and OHISB. Nevertheless, before asserting a lack of association between eHealth literacy and OHISB some contextual influences should be considered, with the first factor relating to the eHEALS tool. Although eHEALS has been repeatedly validated as a reliable measure of eHealth literacy aptitude it did not appear to accurately reflect the online health information seeking skills of all our participants. However, as discussed in Section 3.5, eHEALS is a self-measure of online health information searching skills and it is possible that some of the participants may have overestimated their eHealth literacy skills. This may have contributed to a lack of correlation between their recorded scores and the OHISB characteristics measured. The second critical factor that must be considered relates to the context in which the health information seeking took place. The participants were seeking information during an experiment and were not seeking information for a salient health issue. Thereby some of the participants may have been less inclined to utilise all the strategies that they would employ when seeking information for an emotive health issue, and perhaps did not exhibited their full set of skills for seeking and utilising online health information. Again this may have contributed to the lack of relationships between eHEALS scores and OHISB, and moreover may also have contributed to the weak correlations between the NVS scores and OHISB.

Although bivariate analysis can definitively establish statistically relevant relationships between the derived data, answering the research question requires consideration of other aspects of the participants' OHISB. Furthermore, when considering the OHISB behaviour of the participants the educational background of the study population must also be acknowledged. This was a highly qualified group of participants, many of whom were university graduates, and therefore it would be expected that most of the subjects would have advanced skills to critique and use online health information. Moreover, academic achievement was identified as a significant predictor of both health literacy and health question score in this study. Despite the range in NVS and eHEALS scores, all of the participants exhibited health literacy and eHealth literacy competencies to utilise health resources to accurately answer at least one of the health questions. The most notable behaviour during the study was a reliance on search engines to guide information gathering, a common trait amongst online health information seekers [23]. However, the participants appeared confident to use search engines, formulate search

queries and select suitable resources from results lists, and successfully used this approach to find relevant information to answer the health questions. Another prominent characteristic of the study population was a reliance on unaccredited health information. However, analysis of the types of information resources used indicated an increased tendency to visit and utilise accredited sources of health information. Moreover, evaluation of the manner of accessing results from SERPs also suggests that in many cases accredited sources were selected first, despite not being uppermost on the results list. In conclusion, the results from the study have highlighted a range of strategies that are employed by information seekers when searching for health information online. Although some relationships between health literacy and OHISB were identified, statistically significant relationships between eHealth literacy and OHISB were not found. The significance of the educational background of the study population must also be reiterated, as it is probable that having advanced literacy and analytic skills would also have contributed to the success of the participants. Moreover, it is difficult to accurately assess the extent to which these skills did influence the OHISB of the participants. However, despite mixed results from the statistical analysis, and concerning behavioural traits such as the use of unaccredited health resources, the participants exhibited proficient health literacy and eHealth literacy skills to locate, scrutinise and evaluate health information from a variety of sources, and utilise this information to answer health questions.

3.6.1 Recommendations for Future Research

As indicated in Section 3.5 and 3.6, several challenges were encountered during the study. Therefore, several recommendations for improvement to the study are now provided.

- The most prominent limitation of this study relates to the lack of diversity in the age and educational background of the study population. As described in Section 3.5, it is challenging to assess to what extent these characteristics influenced the behaviour of the subjects during the study. Therefore, a primary recommendation for improving the study would be to ensure that there is greater variation amongst the participants with relation to age and academic achievement. This may reduce the impact of the confounding factors which were present in the current study, and moreover, may also strengthen the statistical validity of the results.

- As highlighted in Section 3.5 there is some uncertainty as to whether the eHEALS tool accurately measured the eHealth literacy skills of the participants. Therefore, in order to improve the validity of the methodology it is recommended that another approach to measuring eHealth literacy be employed. Since the study took place a number of new instruments for measuring eHealth literacy have been proposed. These include the *electronic Health Literacy Scale (e-HLS)* [384], a tool which evaluates skills for using eHealth information, and the *Digital Health Literacy Instrument (DHLI)* [385], which assesses both Health 1.0 and Health 2.0 skills. These tools could be further investigated so as to determine whether a valid and reliable replacement for the eHEALS instrument could be found.
- Throughout the study the HCI Browser proved a reliable tool for recording a user's interactions with a browser, and moreover, was also intuitive for the participants to use. However, the main disadvantage of using this plugin was that each health question was presented alongside a Google search screen, thereby providing a convenient means to start searching for information. One means by which to reduce any bias towards using Google would be to develop a new plugin to collect browser interaction data. However, this plugin would instead present a main web page which relates to the study. This may also encourage the subjects to take longer to consider their strategy for finding health information instead of immediately starting a search for information through the Google search engine.

3.7 Future Work

One notable output from the study was the identification of an association between health literacy skills and health question score. This result is comparable to other research findings which have suggested that those with reduced health literacy skills may experience difficulties when understanding and applying written health information [14, 386]. Studies have suggested that those with inadequate health literacy can be supported by personalised patient education that has been tailored to target various health literacy levels [64, 386]. Thereby the following chapters will investigate whether patient education could be adapted to the health literacy needs of different patients. This will involve the development of a software architecture that provides web-based personalised education for diabetic patients. The education will be adapted to the literacy needs and health status of each patient, and age and gender appropriate images

will also be included in order to enhance engagement with the content. Chapters 4 and 5 will describe the design, implementation and testing of the architecture, and Chapter 6 will describe analysis of engagement with the personalised education.

Chapter 4

Ontological Modelling and Rule-Based Reasoning for the Provision of Personalised Patient Education

4.1 Introduction

Decreasing levels of physical activity and an escalation in unhealthy diets have contributed to an increased international prevalence of diabetes [387]. It is estimated that 415 million people are currently diagnosed as diabetic whilst a further 193 million remain undiagnosed [387]. Diabetes is a chronic condition in which the body may not produce sufficient amounts of the hormone insulin, or may not be able to use the produced insulin effectively [387, 388]. Consequently high levels of glucose can remain in the blood, which can lead to physical health issues [387, 388]. The medical complications of diabetes can include cardiovascular disease, kidney disease, neuropathy, retinopathy and diabetic foot, moreover the condition can have a psychological impact on an individual resulting in increased stress or depression [388, 389]. Complications such as these can reduce an individual's quality of life, and may result in disability or death [387, 390]. Managing diabetes can be a complex and demanding responsibility for the patient [391, 392] and may involve, for example, daily monitoring routines, managing medication, and lifestyle changes such as healthy nutrition and increased exercise [388]. As denoted in Chapter 2 health literacy is an individual's competence to obtain, understand and apply health information [30], and is considered a prominent factor in a patient's ability to make health decisions and to carry out self-management and medication practices [33, 78, 86]. Diabetes self-management routines can be especially challenging for patients with inadequate health literacy skills. For example, studies have identified an association between reduced levels of health literacy and increased risk of hypoglycaemia [393], unintentional nonadherence to medication regimes [43], and reduced diabetes knowledge [394].

One fundamental approach to assisting patients with the management of diabetes is through the provision of high quality education [387, 392, 395, 396]. There is an increasing emphasis that education should enhance a patient's knowledge and skills to manage their condition, and empower them to take an active role in their treatment [390, 392, 395, 397]. Structured educational programs have been successful in assisting

patients with self-management, increasing knowledge, achieving improved HbA1c (Haemoglobin A1c) and blood glucose control, and enhanced quality of life [398-402]. However, although education has proven beneficial for patients with various health literacy competencies [156, 403] there is a concern that diabetic patient education is not always accessible for patients with limited health literacy skills. For example, a review of printed diabetes brochures by Hill-Briggs and Smith [404] found variable levels of readability in the brochures, and proposed that not all the content would be understandable for patients with lower health literacy. It has been suggested that diabetic education will be most effective if it is individualised to the particular medical, educational, social and health literacy needs of each individual [405, 406]. Moreover diabetic education should be an ongoing process that adapts to the changing health needs of a patient [395]. Diabetic patients themselves have expressed a preference for timely, personalised information, whilst also indicating that excessive amounts of information can be overwhelming [405].

Personalisation presents a means to overcome the concerns associated with generic patient education by enabling the provision of education that focuses on the particular needs and health objectives tailored for each patient. Personalisation has been used in health communication to motivate patients to process health messages and improve their health behaviours [165, 176]. Personalised education may be more accessible, usable and engaging for the patient by providing advice that is more relevant to their lifestyle when compared to generic educational approaches. In order to investigate whether patient education could be adapted to the individual needs of each patient, it was decided to develop a web-based architecture that could tailor education for diabetic patients. However, providing personalised patient education requires access to, (1) accurate and up-to-date information about the patient and their health status, and (2) clinical information related to the health domain. Ontologies can provide the expressiveness to represent semantics for a range of data, and have been widely utilised to capture user profiles and domain knowledge for a range of healthcare applications [303, 318, 319, 407]. Thereby it was decided that an ontology would be included in the architecture to represent a model of the domain knowledge, and that semantic rules would provide reasoning for the personalisation mechanisms. Thereby the second research question of this research study is proposed.

- Can a semantic web ontology and rule set support the personalisation of patient education as suitable for the health information needs and health literacy of an individual?

This chapter describes a study which was designed to investigate the research question. The aim of the study was to investigate whether semantic web technologies could be used in a web-based architecture to facilitate the creation of personalised diabetic education. This chapter describes the design, development and testing of the architecture. A central component of the architecture comprises an OWL ontology which represents the domain knowledge necessary for the production of the personalised education. The ontology includes a user model that captures information about the patient, and a model of the health condition diabetes including symptoms, treatments and complications. Semantic Web Rule Language (SWRL) [317] rules and a Pellet [408] reasoning engine provide decision support for the personalisation mechanisms. The architecture provides educational content that is tailored to focus on the particular health objectives and personal characteristics of each patient, and is transcribed at a readability level suitable for the patient's health literacy needs.

The remainder of the chapter will describe the methodology by which the components of the architecture were designed, developed and tested. Section 4.2 reviews related work in the sphere of web-based diabetic patient education, and also examines how semantic web technologies have been used in other diabetic healthcare applications.

4.2 Related Work

4.2.1 Web-Based Applications for Diabetic Patients

Current approaches to structured diabetes education include face-to-face or group based education scenarios [397]. However, it has been suggested that eHealth applications could be used to deliver diabetes education and provide assistance with healthcare [397, 409]. The use of electronic media to provide education may reduce the financial costs associated with current delivery formats, and moreover could enhance convenience for the patient as the education could be accessed at an expedient time and location [397, 410, 411]. The effectiveness of web-based diabetic education is often evaluated by quantifying outcomes such as patient learning, behavioural changes or changes in HbA1c levels. However, there has been inconsistency in the results [412]. For example,

a web-based self-management program that included weekly informational sessions, nutritional support and glucose monitoring, reported longer term (18 months) improvements in self-efficacy and activation for participants [413]. A similar study [414] which compared health outcomes for subjects using a computer-assisted diabetes self-management intervention, and a group that used an intervention with less procedures, found improved physical activity, nutritional habits and fat intake behaviours amongst the computer-assisted intervention group. However, the results also indicated that the effectiveness of the intervention was most significant during the first four months of the study, and that effect size was modest at 12 months. It has been suggested that automated tailoring would enhance the effectiveness of diabetic self-management interventions [414].

This study focuses on the development of web-based personalised diabetic education. A significant aspect of the study is a focus on adapting education provision to the health literacy needs of diabetic patients. A review of associated literature indicated that although there is a rich body of research related to the development of electronic applications for diabetic education, there are fewer applications that focus on providing health literacy appropriate solutions. Kandula et al. [156] developed a multimedia diabetes education program that was suitable for various health literacy competencies. The use of plain language, images, definitions and layout were formulated to encourage learning and memorisation. Moreover, patients were also involved during the design stages to ensure that the materials were comprehensible and appealing. Analysis of the system with 190 patients indicated that diabetes knowledge improved significantly for patients with inadequate and adequate health literacy, although the knowledge improvement was greater for the latter group. A multimedia application which was targeted at lower literacy individuals is described by Gerber et al. [158]. The application incorporated audio and video in order to reduce text information, and was provided on touch-screen kiosks for ease of navigation. Interventions were tailored in accordance with computer literacy and learning style. However, although the results indicated an increased awareness of perceived susceptibility to diabetes complications amongst the subjects, no significant differences were found in relation to diabetes knowledge, self-efficacy or A1c. Personalised features were incorporated into an online disease management system which assisted diabetic patients with self-management of their condition [415]. Patients could directly enter their health data or use clinical devices

that would upload glucose readings. Video or text information that was personalised in accordance with the patient's health status was then provided. A six-month comparison of subjects using the system (intervention group) and a usual care group indicated that the intervention group had achieved significant reductions in A1c. However, at 12 months there was no significant difference in A1c control between the two groups.

The reviewed literature suggests that benefits can be derived from the use of web-based diabetic education. However, there is variation in the effectiveness of the education as measured by various health outcomes. Adaptation of education to the health literacy needs of patients has been examined as a means of enhancing the usability of diabetic patient education. However, the lack of related studies suggests that further research is needed to identify useful computerised methods for tailoring education to health literacy competencies, and to appraise the efficacy of these approaches.

4.2.2 Ontological Models of Diabetes

An ontology has been defined as “*a shared understanding of some domain of interest*” [259]. It reflects a particular world view of a domain by representing domain knowledge as an arrangement of concepts, definitions and relationships, and defining shared agreement on the meaning of these components [259]. Modelling data in an ontology ensures that the meaning of the data is unambiguous and is therefore suitable for processing by reasoning technologies [256, 416]. Ontologies are being increasingly used within the biomedical and health domains to standardised medical vocabularies and provide validated semantics for medical data [268]. Within the domain of diabetes, ontological models have been used to express a range of clinical knowledge related to diabetes management. Arwan et al. [417] proposed the use of ontologies to facilitate the production of food recommendations for diabetic patients. The proposed architecture incorporates a Calorie Foods ontology which is used to represent various characteristics of diabetic patients including calorie needs. A Foods ontology is used to represent food attributes. SPARQL Protocol and RDF Query Language (SPARQL) queries are used to derive semantic matches between the ontologies and produce the nutritional recommendations. An ontology based recommendation system that supports physicians in prescribing diabetic medication is presented by Chen et al. [418]. An ontology is used to capture medication related knowledge such as medication name, side effects, contraindications and associated monitoring. Medical tests such as HbA1c, liver tests

and renal insufficiency tests are also modelled, and are utilised during reasoning processes to determine suitable medication for a particular patient. The suggestions provided to the physician include information related to dosage, contraindication, side effects and the monitoring required for a recommended drug. Chammas et al. [419] proposed a tool for diabetic patients that provides graduated levels of advice for avoiding diabetic podiatry related complications. Central to this is a computational model that comprises an ontology and semantic rules. The ontology captures patient information such as podiatry observations, symptoms, lifestyle factors and medical test results. Semantic rules then determine the category of advice provided to a patient, for example a patient with tenderness in their feet would receive advice that encourages a visit to a footwear specialist. The rules can also filter out advice that is irrelevant to an individual, for example the system would not send a reminder about foot hygiene to a patient that regularly attends to this aspect of foot management.

Ontologies have also been developed that capture the knowledge necessary to assist with the identification of diabetes. Rahimi et al. [420] developed the Diabetes Mellitus Ontology (DMO) to assist with the diagnosis and management of individuals with diabetes. An algorithm was developed, based on the DMO, which utilised data derived from electronic health records to identify patients with type 2 diabetes mellitus. Alharbi et al. [421] developed an ontology based clinical decision support system to diagnose diabetes and provide an appropriate treatment plan. Clinical practice guidelines were utilised to develop a diabetes domain ontology, a patient ontology and SWRL diagnostic rules. The rules utilise data captured in the ontology, including laboratory tests and symptoms, to identify cases of diabetes and pre-diabetes.

The reviewed literature highlights how ontologies can be used to provide a concise model of the medical, contextual and patient information associated with diabetes diagnosis and treatments. Semantic reasoning is often used to facilitate decision making capabilities. Thereby ontology based applications can facilitate the provision of individualised healthcare and treatment for diabetic patients, which may prove useful for both physicians and patients. In common with much of the literature reviewed, the architecture presented in this chapter uses an ontology and rule-base to personalise health information. However, the main novelty in the architecture is the range of personalisation provided which concentrates on both content and presentation. The

education is tailored to the specific health needs of each patient, and provides information on a range of pertinent patient experiences such as symptoms, complications and treatments. In order to assist the patient's comprehension, the text information is adapted to an appropriate readability level. Moreover, readability is enhanced by displaying the text at an appropriate size and style. Engagement is encouraged by including images that have been personalised by age group and gender. Furthermore, the education will also adapt to changes in the patient's health status or contextual factors such as lifestyle.

4.3 A Web-Based Architecture for Personalised Patient Education

This chapter describes the design and development of a web-based architecture that creates personalised education for diabetic patients. The architectural model, which is illustrated in Fig. 4.1, comprises (1) a Spring MVC [422] controller which handles communication between the client and the server, (2) logic components, (3) an ontological knowledge base, (4) rule-based personalisation and contextualisation mechanisms, and (5) a database. The education is constructed as a JavaServer Page (JSP) [423]. A layered architecture was designed in order to separate the knowledge base from the personalisation mechanisms and other functional components.

The *Modelling and Management layer* contains the knowledge base, which is modelled as an OWL ontology. The *patient education ontology* represents information relating to the four main domain entities, (1) the patient, (2) the medical condition diabetes, (3) physical activities, and (4) the educational content. As denoted in Fig. 4.1, each entity is represented as an ontological model which comprises hierarchies of classes, and class properties to define relationships with other classes and data values. In this way, the model can reflect how information is organised for this entity in the patient education domain. The four main models included in the ontology are as follows.

- The *Patient Model* represents information related to the patient which is essential for personalisation of the education. This includes personal information such as age and gender, educational information, and details of the patient's health status.

- The *Medical Conditions Model* contains information associated with the health condition diabetes, including symptoms, treatments and complications.
- The *Activity Model* contains information related to physical activities such as associated intensity and location of the activity.
- The *Educational Content Model* includes details associated with the text and image educational components. This may include, for example, the readability level of a piece of text.

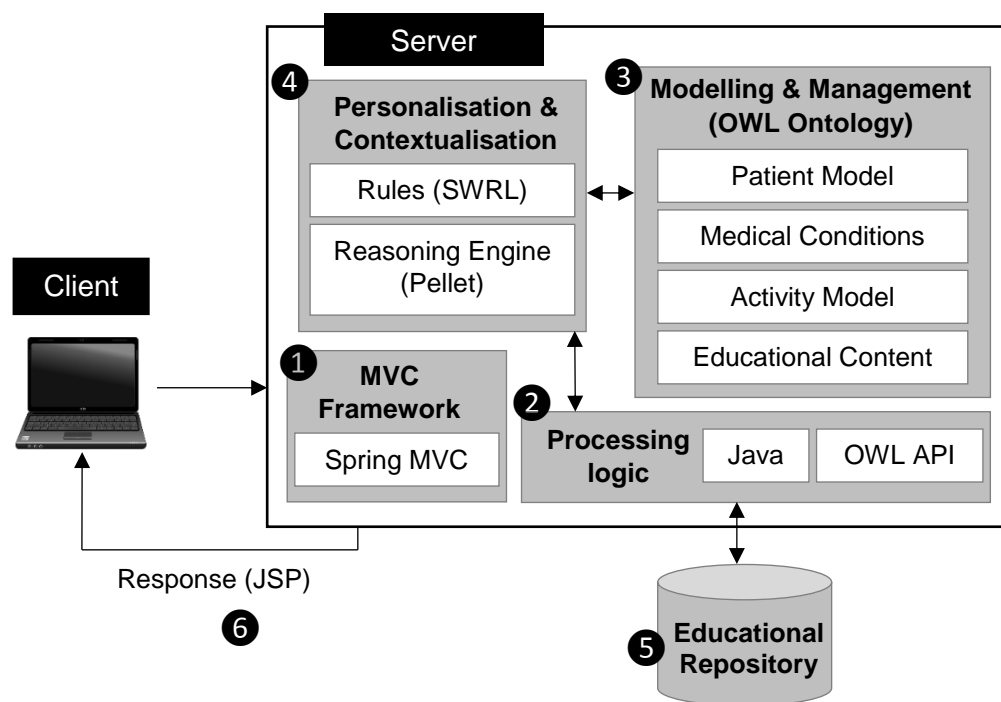


Fig. 4.1. Architecture to provide web-based personalised patient education. (1) Spring MVC handles communication with the client, (2) processing logic handles interactions between server-side components, (3) domain knowledge is represented in an ontology, (4) a reasoning engine is utilised to execute personalisation rules, (5) suitable content is retrieved from the educational repository, and (6) delivered to the patient as a web page.

The *Personalisation and Contextualisation* layer contains a set of SWRL personalisation rules and a Pellet reasoning engine. The rules utilise the knowledge in the ontology to determine the composition and style of the education. The educational

content, which includes text and image components, is maintained in the *Educational Repository*. The personalisation rules determine the combination of text and image components that are selected for each patient. The educational content is formatted as a web page and presented to the patient through a web browser.

4.4 An Ontological Model for Personalised Diabetic Patient Education

4.4.1 A Methodology for Ontology Design

At the beginning of the project several ontologies were reviewed (through the OBO Foundry website [286]) in order to determine whether using a larger domain ontology would be a suitable approach. However, it was decided to author a novel application ontology. The main reason for this was that a domain ontology would contain content that was superfluous to the requirements of the patient education architecture and thereby sections of the ontology would remain unused. Moreover, there was a concern that reasoning on a larger domain ontology may be computationally expensive [424]. Engineering an ontology specifically for the patient education architecture also ensured that all the data and relationships essential for the tailoring of patient education were present in the ontology. Moreover, creating a smaller and less complex ontology also assisted with making changes and updates to the ontology during the initial phases of development.

Fig. 4.2 indicates the steps involved in the development and evaluation of the patient education ontology. The design and implementation of the ontology will be detailed in this chapter, and evaluation will be discussed in Chapter 5. Some of the activities involved in the design and construction of the ontology were based on the modelling approaches described in [261] and [300], however the described activities were modified as required to develop a model of the patient education domain.

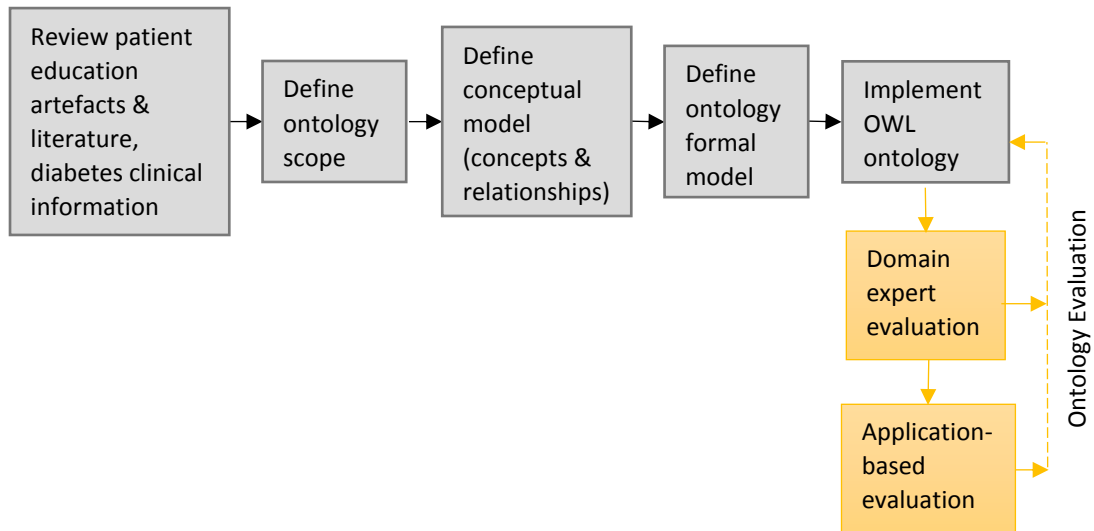


Fig. 4.2. Stages of engineering for the patient education ontology. Design and implementation phases are grey and evaluation stages are yellow.

Review Artefacts and Literature

The key aim of this project was to provide personalised diabetic patient education, therefore it was essential to comprehend, (1) current approaches to diabetic patient education, and (2) the factors that limit the effectiveness of generic patient education. The initial design activities focused on acquisition of this knowledge. The primary design stages included a review of the current approaches to diabetic patient education including pamphlets, booklets and educational websites from government and accredited health organisations. Some of the information sources used included Diabetes UK [388], the American Diabetes Association [389], the UK National Health Service (NHS) [66], the International Diabetes Federation (IDF) [425], and the U.S. National Library of Medicine [68]. This facilitated the identification of key medical topics that were essential for inclusion in diabetic education. A literature review which focused on generic patient education was also conducted. This helped to identify a number of factors that could limit the effectiveness of generic educational approaches. The following *objectives* were defined for the personalisation of the patient education.

- The education would be tailored to the health status of each patient, and include information for their particular experience of symptoms, treatments and

complications. Information that is not pertinent to the patient's health status would be excluded.

- In order to assist the patient's comprehension, the text information would be adapted to a suitable readability level.
- Readability would be enhanced by displaying the text at a size and style suitable to the patient's needs or preferences.
- In order to encourage engagement, images would be personalised by age group and gender.
- The layout of the education would be suitable to the patient's experience of using web-based applications.

Define Ontology Scope

The initial reviews assisted with defining the scope of the ontology and the domains that would be represented. The ontology would represent the domain of diabetes and would include a range of symptoms, complications and treatments. As physical activity is a major aspect of diabetes management it was decided that this domain should also be modelled in the ontology. It was also determined that the personalisation mechanisms would focus on tailoring the content and presentation of the education to the health status and personal characteristics of the patient, and that a user model would be designed accordingly. In addition, aspects of the educational content would also have to be represented, which would include attributes of the text and image components.

Define Conceptual Model

The following stage involved defining a conceptual model of the ontology. Information was obtained from the sources used during the initial design phase [66, 68, 388, 389, 425], and was utilised to develop a knowledge base of the clinical features related to the symptoms, complications and treatments of diabetes. In order to design a conceptual model, concepts were extracted from the knowledge base and the relationships between these concepts were determined.

Define Formal Model

In the next stage a formal model of the ontology was defined. This involved organising the classes into hierarchical relationships. A top-down design approach was used whereby the classes in the top layer of the ontology define models of the four main domain entities, (1) the patient, (2) the health condition diabetes, (3) physical activities, and (4) the educational content. The information captured for each model becomes increasingly specialised with each subsequent layer. The properties of each class were defined, and axioms were assigned to formalise the relationships between class and data properties. When the formal model had been finalised, the ontology editor Protégé [426] was used to implement the OWL ontology. The Pellet reasoner was also used to check the correctness and consistency of the ontological model. The evaluation phases involved both expert appraisal and application-based evaluation, and will be discussed in Chapter 5.

4.4.2 User Modelling

Personalising diabetic patient education would require various information to be held about the patient including their personal characteristics and preferences, health status, physical activity capabilities, and educational competencies. Thereby a *user model* was designed to capture pertinent data about the patient. In the ontology the user model is represented as the *Patient Model*. The objectives for personalising the education had been established during the initial literature review (as described in Section 4.4.1). The Patient Model was designed to hold user data that would be needed to facilitate tailoring the content and presentation of the education to the health information needs and preferences of a patient. The initial design phase focused on identifying clinical data that was pertinent to diabetic patients such as diabetes health checks. Various health information resources [66, 388, 389, 425] were reviewed and relevant clinical attributes were identified including blood glucose level, height, weight, body mass index (BMI) and possible co-morbidities. These characteristics would be used to determine the patient's health status including, for example, whether a patient was overweight, and subsequently whether the education should include information related to healthy nutrition. Furthermore, the Patient Model would also represent the symptoms and complications experienced by a patient and any prescribed treatments, as much of the educational content would focus on these. It was also decided to model a patient's highest qualification and preferred language to facilitate tailoring the readability of the

text content. The patient's name would also be required for inclusion in the content so as to enhance their identification with the education. As the educational images would be personalised by age group and gender, personal characteristics including age, gender and ethnicity were also added to the Patient Model. The patient's level of eHealth literacy and preferred colour palette were also modelled so as to adapt the presentation of the education to their technical ability and aesthetic preferences.

The personalised patient education would include information related to physical activity. Therefore a review of sports and activity information [66, 427] was undertaken in order to establish which patient attributes would be needed to provide advice suitable for the patient's fitness level and interests. It was decided that a patient's preferred activities and activity environment (indoors or outdoors), and physical activity level would be recorded. Moreover, the patient's home and work address would also be stored as this information was required to contextualise the physical activity advice as suitable for the patient's lifestyle.

With such a complexity of information to be captured it was determined that the Patient Model would be categorised as seven main profile classes. The data held in each profile class is detailed in Table 4.1, and Fig. 4.3 provides an illustration of several of the main Patient Model classes and properties in Protégé. Fig. 4.4 indicates the organisation of the profile classes within the Patient Model, denoting the main asserted data and object properties within each class. As indicated in Fig. 4.4, several of the profile classes have relationships with classes from the Medical Conditions Model and Activity Model. Section 4.4.3 will describe how relationships between the Patient Model and the other entity models enable the creation of the personalised education.

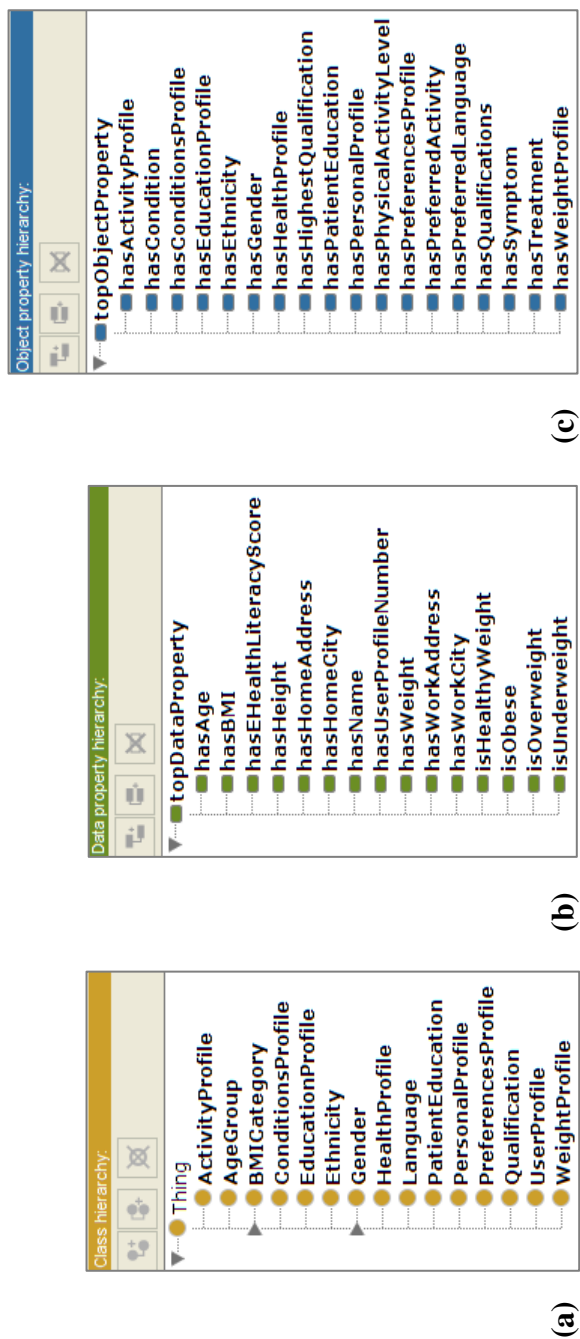


Fig. 4.3. An excerpt of the Patient Model illustrating (a) classes, (b) datatype properties, and (c) object properties.

Table 4.1. An excerpt of the key profile classes in the Patient Model.

Class	Data modelled
ActivityProfile	This class describes a patient's activity objectives. This includes their current level of physical activity, their preferred activities and preferred activity environment (indoors or outdoors). This information will be used to select physical activities that are suitable for the patient.
EducationProfile	This class describes a patient's educational background. This includes their highest qualification. This information will be used to determine which readability level is the most appropriate for the text components included in the patient's education.
PersonalProfile	This class describes a patient's personal characteristics including their age, gender, ethnicity, eHealth literacy capabilities, and work and home address.
PreferencesProfile	This class describes a patient's preferences for the composition and layout of the education. This includes their preferred language for text components.
HealthProfile	This class organises information related to the patient's health.
ConditionsProfile	This class describes clinical data related to the patient's health conditions. This includes co-morbidities, symptoms, treatments and complications. This information will be used to select educational components that relate to the aspects of diabetes that are pertinent for the patient.
WeightProfile	This class describes clinical data related to the patient's health status. This includes weight, height and BMI. This is used to determine which educational components should be included.

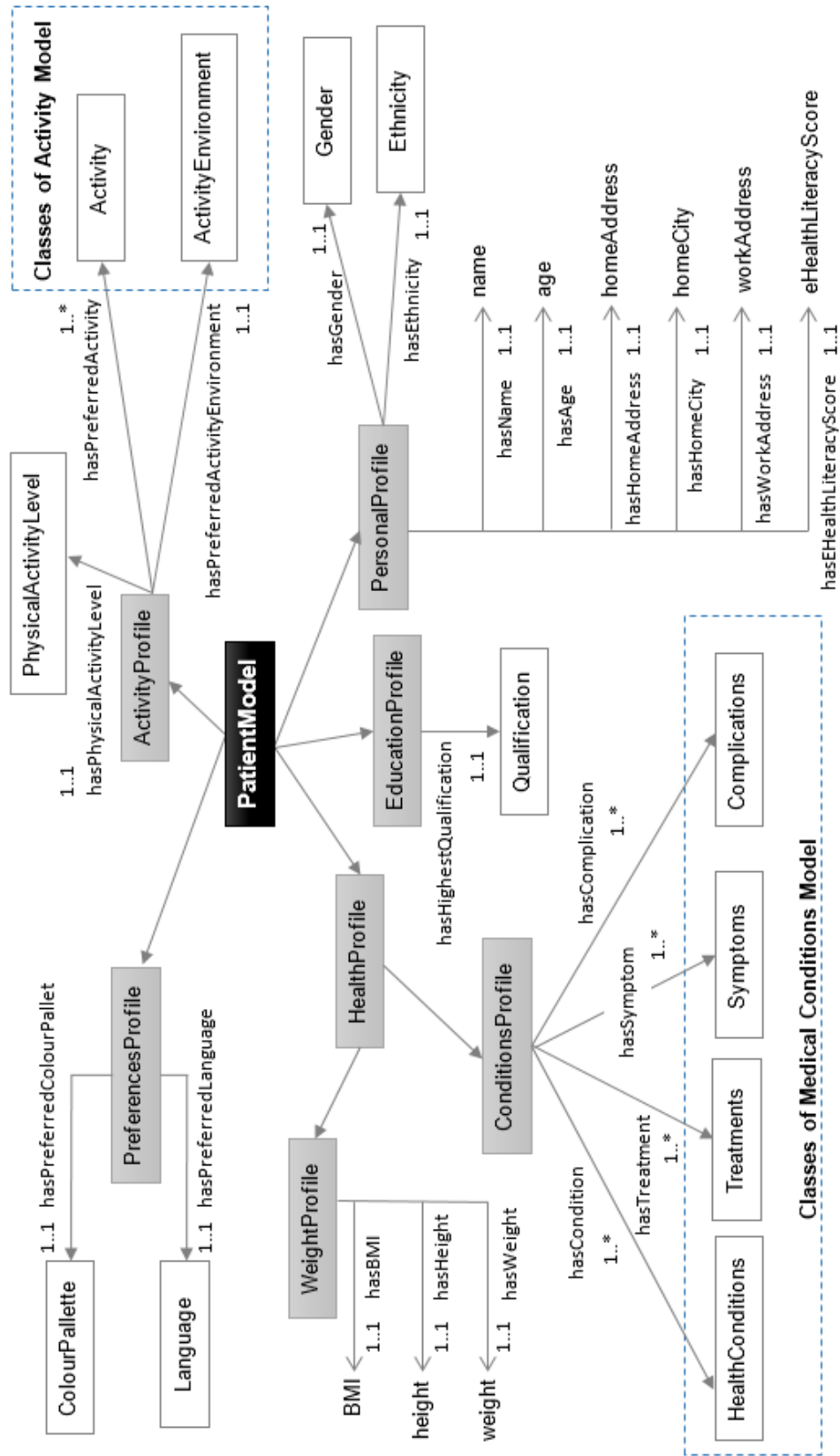


Fig. 4.4. Overview of the Patient Model, indicating main profile classes (grey) and associated asserted datatype (forked arrows) and object (boxes) properties.

4.4.3 Associations Between the Ontological Models

Fig. 4.5 indicates how associations can be inferred between classes represented in the Patient Model, the Educational Content Model and the Activity Model. The classes in the top layer of the ontology (black) define models of entities within the domain and the next layer (dark grey) classifies the information that will be captured in each model. Personalisation is enabled by capturing characteristics of the patient, the educational content and physical activities as property values (light grey) in lower level classes. Associations are derived from the properties of each class and are utilised to adapt the composition and presentation of the education to the preferences and needs of each patient.

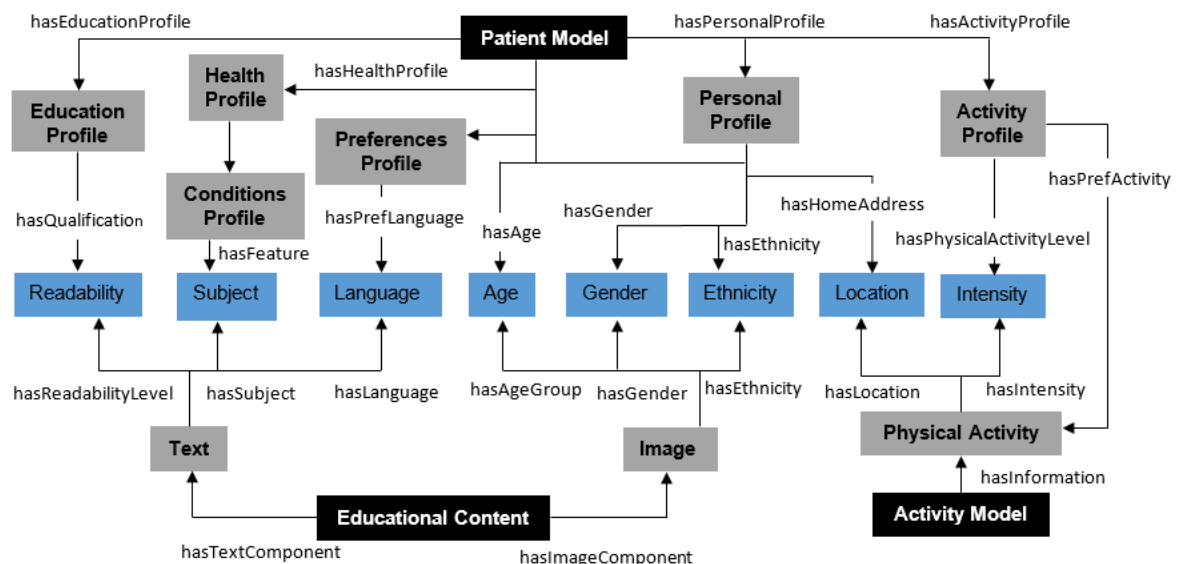


Fig. 4.5. Association characteristics of the Patient Model, Educational Content Model and Activity Model.

The Educational Content Model

The educational content comprises text and image components such as pictures and diagrams. The Educational Content Model contains three classes, EducationalContent, Text and Image. The properties and relationships modelled can be used to infer associations with properties that are captured in the Patient Model. In order to enhance a patient's identification with the education it will contain images that are personalised to their age group, gender and ethnicity. Image properties including hasAgeGroup, hasGender and hasEthnicity provide a means to infer associations

with similar properties captured about the patient. This will ensure that images that are highly correlated with the patient's profile are selected from the education repository. Class `Text` represents text information that pertains to aspects of diabetes. For example, a paragraph of text could describe a particular symptom of diabetes. Text components are categorised by several properties including `hasLanguage`, `hasSubject` and `hasReadabilityLevel`. These properties can be associated with properties of the patient's `EducationProfile` and `PreferencesProfile` to ensure that the text information matches the readability needs and language preferences of the patient.

The Medical Conditions Model

The Medical Conditions Model represents clinical features of diabetes and is categorised into three strands; symptoms, treatments and complications. Each strand comprises subsumption relationships that provide increasingly specific representations of clinical data. For example, the data captured in the treatments model is categorised as pharmacological and non-pharmacological treatments, and in the subsequent layer pharmacological treatments is further classified as diabetic medications and insulin types. Representing the knowledge with increasing granularity enables an extensive range of diabetes related clinical data to be captured. Fig. 4.6 illustrates several of the subsumption relationships within this model. Although aspects of diabetes are common to many patients it is likely that not all patients will have the same combination of symptoms, health concerns and treatments. Personalisation can ensure that the education provided focuses on the particular aspects of the condition that are relevant to each patient. Data regarding a patient's health status is captured in the Patient Model class `ConditionsProfile`. The patient's combination of symptoms, complications and treatments are represented through relationships with instances of the Medical Conditions Model (as illustrated in Fig. 4.4). This information can be utilised to ensure that the education components selected are targeted to the pertinent aspects of the condition. For example, for a patient with type 2 diabetes, the information provided could relate to treatments such as cholesterol checks and weight management, while for another patient with the same condition the education may concentrate more on diabetic medications. Tailoring the education in this way ensures that the education is focused on the needs of the patient and that it changes dynamically in response to changes in

their medical status. Furthermore, the patient is not overwhelmed by information that is not directly applicable to their needs.

The Activity Model

One aspect of treatment for diabetes relates to increased physical activity. The Activity Model was designed to model various aspects of physical activity. Each patient has an activity profile which denotes their preferred activities and current physical activity level. Associations between these two models can be utilised to ensure that activity suggestions are focused on the patient's preferences and are suitable to their current physical capability. Nevertheless, while it is favourable to provide information that is personalised to the preferences of the patient it must also be contextualised to their lifestyle. One means by which to achieve this is to consider whether it is convenient for the patient to take part in this activity. Location related information is captured for activities and can be used to determine whether this activity takes place within a specified distance of the home or workplace of the patient. This ensures that the activity advice provided is relevant to the patient's lifestyle and is applicable to the different environments in which they reside.

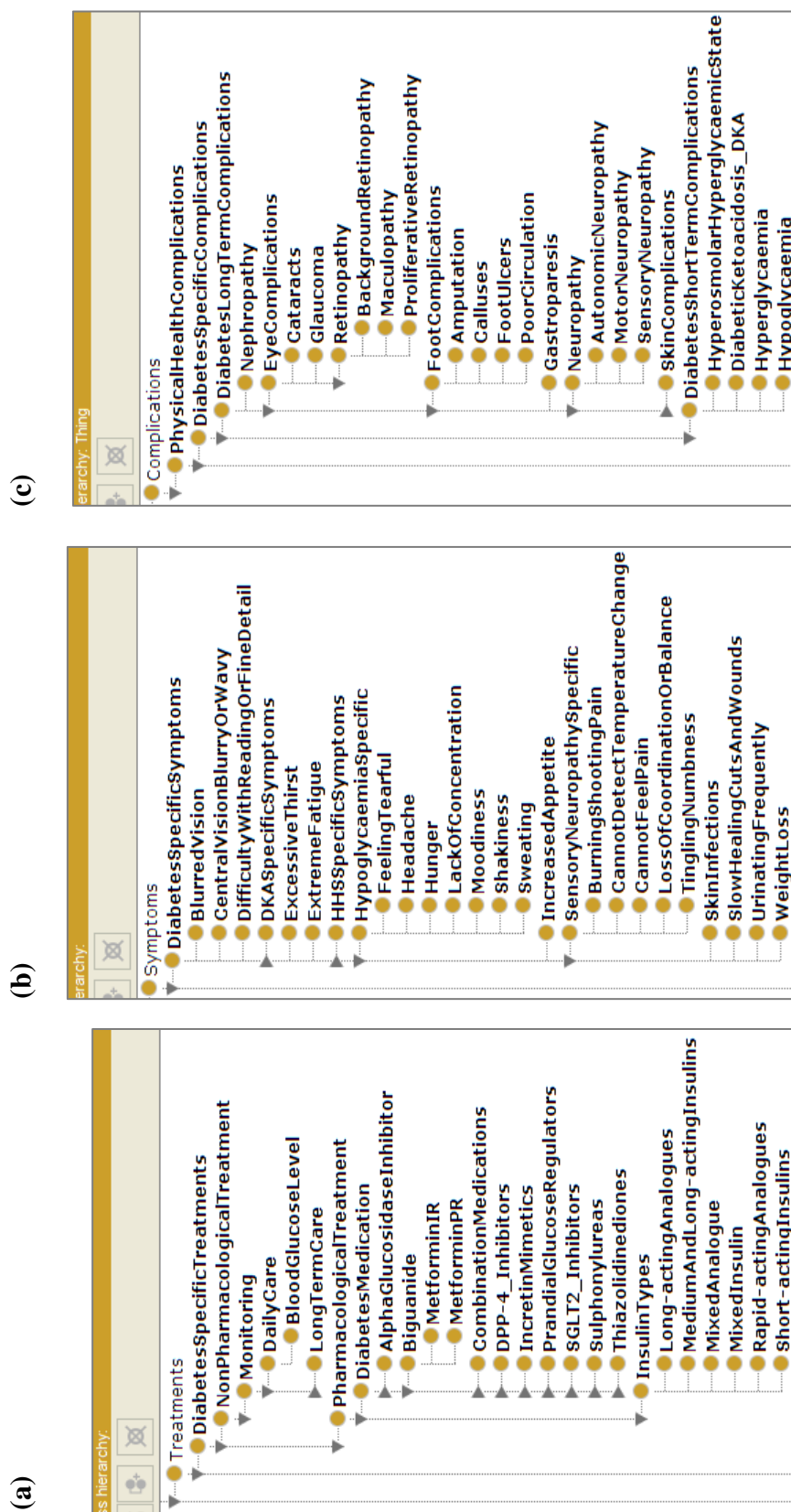


Fig. 4.6. Subsumption relationships within the Medical Conditions Model (a) treatments, (b) symptoms, and (c) complications.

4.5 Rule-Based Personalisation

The models captured in the ontology provide templates for representing knowledge about the domain entities. Classes and properties express the data that should be captured for each entity and how it should be organised. In order to store the information corresponding to a specific entity, an individual member of a class must be created. An *individual* is an instance of a class in which the properties have been filled with data values. The ontology is populated with multiple individuals to represent specific instances of domain entities. The *Personalisation and Contextualisation* layer contains a set of rules that utilise the information captured in these individuals to determine the composition, style and readability of the education. The rules are represented using the Semantic Web Rule Language [317]. SWRL is a rule language that is built on OWL, and provides additional expressiveness to capture knowledge that cannot be expressed in OWL. SWRL rules will reason using the individuals in the ontology and deduce new knowledge that is added to the ontology.

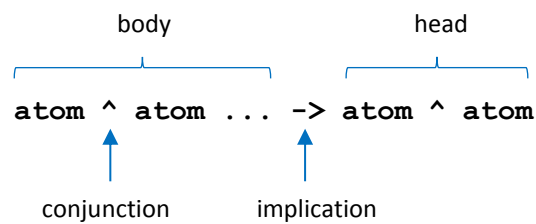


Fig. 4.7. Structure of a SWRL rule [428].

Fig. 4.7 denotes the syntax of a SWRL rule. The rule comprises an *antecedent* component (body) and a *consequence* component (head), which are separated by the operator \rightarrow . The antecedent and consequence contain positive conjunctions of atoms which describe the conditions to be evaluated, and a rule specifies that if the atoms in the antecedent are true then the atoms in the consequence part are also true [428]. SWRL rules are expressed by using OWL concepts. For example, the first rule in Table 4.2 uses references to named classes, properties, built-ins and variables in order to assign a patient's BMI category. The first atom defines that a patient's weight profile is a member of class `WeightProfile` and is represented by the variable `?prof`. This individual has a datatype property, `hasBMI`, which indicates the patient's BMI score. The built-in `greaterThanOrEqual` is used to determine if this score is within the range

of the *obese* category. If all the atoms in the antecedent evaluate to true then the atom in the consequence must also evaluate to true, thereby the boolean property `isObese` is inferred for `?prof` with the value of `true`.

The rules enable personalisation by asserting conditions that will be evaluated during the generation of the education. For example, a patient's treatments may be appraised in order to filter out unnecessary information about treatments that have not been prescribed. In order to enable the rules to implement personalisation a rule-based reasoner is required. In this architecture Pellet is used to reason with the rules and determine logical inferences about the data captured in the ontology. Inferences made will include new associations between the data and new classifications for individuals. The asserted facts will be added to the ontology and utilised to tailor the composition and presentation of the education to the needs of each patient.

Table 4.2 contains a subset of the rules used in this architecture. Rules 2 and 3 instantiate properties that are evaluated in subsequent rules. For example, Rule 2 will set an age group property for a patient's personal profile. This value is necessary for selecting age relevant images from the educational repository. Similarly Rule 3 exemplifies a subset of rules that will examine a patient's symptoms and complications in order to determine whether the patient will require information about a further complication. In this rule the patient's symptoms are scrutinised and it is determined that they are experiencing Diabetic Maculopathy. This new knowledge is added to their health conditions profile. The role of Rule 4 is to ensure that an erroneous state has not occurred in the knowledge base. This rule inspects a patient's condition profile to ensure that they have not been incorrectly represented as having both type 1 and type 2 diabetes. If this situation has occurred the corresponding individual is classified as a member of class `InvalidDiabetesCondition` to indicate that it contains erroneous data and should not be included in the personalisation processes.

Tailoring the Readability of the Education

In order to ensure that the text information is comprehensible for the patient it is tailored according to their level of health literacy. A patient's comprehension of written health information can be affected by the readability level of the text content, however health

literacy may also be affected by contextual factors. For example, a patient with a sufficient level of literacy may find it difficult to interpret information that contains unfamiliar terms and diagrams [86] or they may find it more demanding to understand information at different stages of their treatment or at times of stress [78]. Rule 5 illustrates how an association can be established between a patient's level of health literacy and the readability level of text components. In this architecture the Simple Measure Of Gobbledygook (SMOG) [429] readability measure is used to classify text components. A SMOG score can be correlated with the UK National Adult Literacy Standards [430], whereby, for example, if a piece of text has a score of 11 or 12 it is assumed that this text is understandable for adults with level 1 literacy skills. The literacy standards can be mapped to the National Qualifications Framework [431], hence in this architecture a patient's highest qualification is used to determine the most appropriate SMOG level for their understanding. For example, Rule 5 evaluates whether a patient's highest qualification is a National Vocation Qualification Level 1, which would correlate with a SMOG score of 11 or 12. Nevertheless, as previously indicated, a patient's ability to understand health information can be affected if they are anxious, unwell or fatigued. These health conditions can be captured in the patient's conditions profile and utilised to adapt the readability level of the text accordingly. Rule 5 assesses whether the patient is experiencing stress and consequently modifies the readability score to 9 or 10.

Tailoring the Content of the Education

Characteristics of the personalised education are captured in the class `PatientEducation`. Each patient represented in the ontology will have an individual of this class which contains details of their personalised education. This class has a number of properties that denote attributes of the education such as the images that will be included and the text style. Personalisation rules are used to infer the values of these properties. The property `hasEducationTopic` indicates the text components that will be included in the education. This may include generic information related to diabetes and material that is focused on the set of symptoms, treatment and complications relevant to the patient. Rules 6, 7, 8 and 9 are a subset of the rules that will identify the topics to be included. Fundamental to diabetes management is blood glucose monitoring, therefore this will form a central component in education for all diabetic patients. Rule 6 ensures that all patients with type 1 diabetes receive information on this

subject. A similar rule for patients with type 2 diabetes also exists. Rule 7 utilises a patient's inferred BMI category to determine whether nutritional information should be included. In this example the consequence will evaluate to true if the patient is obese. Rule 8 represents a subset of rules that ensure that information related to the patient's symptoms, treatments and complications is also included. In particular, this rule ensures that information related to a patient's symptom is added to the education. The rule also ensures that the text content is at an appropriate readability level for the patient and is transcribed in their preferred language. Rule 9 is an example from a set of rules that determine which images are added to the education. This rule selects images that correspond to the physical activities that will be recommended to the patient as part of the education. The images chosen will also be appropriate to the patient's gender, ethnicity and age group.

Tailoring the Presentation of the Education

Rules 10 and 11 focus on adapting the presentation of the education to the needs of the patient. Rule 10 assesses whether a patient is experiencing blurred eyesight, a common symptom for diabetic patients. If this constraint evaluates to true then property `hasTextSize` is inferred to be large. This rule directly assesses the presence of blurred eyesight, however, additional rules exist that will evaluate the presence of other eyesight related complications or the presence of an unrelated visual impairment. Rule 11 utilises a patient's level of eHealth literacy to determine the layout of the education. A patient with reduced eHealth literacy skills may have less experience of using web-based platforms, thereby the presentation of the education can be adapted to be less cognitively demanding for this patient. Rule 11 prescribes that for a patient with an eHealth literacy score of less than 20 a less complex presentation template is assigned.

Tailoring Physical Activity Suggestions

If a patient is prescribed to increase their physical activity as a means to manage their condition then physical activity advice will be included in the education. Information about specific physical activities is captured as individuals in the ontology. The information captured for each activity includes the location where it takes place and an intensity level. Rule 12 illustrates how a match can be established between a patient's physical activity level and a preferred activity with an appropriate intensity level. In order to ensure that the activity is convenient for the patient's lifestyle the rule also

evaluates whether the activity takes place in the patient's home town or in a nearby geolocation. If all the atoms in the antecedent are true then the `Activity` individual is added to the patient's activity profile, thereby ensuring that the corresponding activity information will be provided to the patient in the education.

Table 4.2. Subset of SWRL rules used in Personalisation and Contextualisation layer.

	SWRL Rule
1	WeightProfile(?prof), hasBMI(?prof, ?bmi), greaterThanOrEqualTo(?bmi, 30.0) -> isObese(?prof, true)
2	AgeGroup(?group), PersonalProfile(?prof), hasAge(?prof, ?age), hasMaxAge(?group, ?maxAge), hasMinAge(?group, ?minAge), greaterThanOrEqualTo(?age, ?minAge), lessThanOrEqualTo(?age, ?maxAge) -> hasAgeGroup(?prof, ?group)
3	ConditionsProfile(?conProf), hasSymptom(?conProf, CentralVisionBlurryOrWavy), hasSymptom(?conProf, DifficultyReadingOrFineDetail) -> hasComplication(?conProf, DiabeticMaculopathy)
4	ConditionsProfile(?prof), hasCondition(?prof, DiabetesType1), hasCondition(?prof, DiabetesType2) -> InvalidDiabetesCondition(?prof)
5	ConditionsProfile(?cProf), hasCondition(?cProf, Stress), EducationProfile(?eduProf), hasHighestQualification(?eduProf, NVQLevel1) -> hasLiteracyLevel(?eduProf, Readability9_10)
6	UserProfile(?uProf), ConditionsProfile(?cProf), HealthProfile(?hProf), PatientEducation(?edu), hasHealthProfile(?uProf, ?hProf), hasConditionsProfile(?hProf, ?cProf), hasPatientEducation(?uProf, ?edu), hasCondition(?cProf, DiabetesType1) -> hasEducationTopic(?edu, BloodGlucoseMonitoring)
7	UserProfile(?uProf), WeightProfile(?wProf), HealthProfile(?hProf), PatientEducation(?edu), hasHealthProfile(?uProf, ?hProf), hasWeightProfile(?hProf, ?wProf), isObese(?wProf, true), hasPatientEducation(?uProf, ?edu) -> hasEducationTopic(?edu, HealthyEatingAdvice)

8	<p>ConditionsProfile(?cProf), HealthProfile(?hProf), PatientEducation(?edu), PreferencesProfile(?prefProf), Symptoms(?sym), TextContent(?txtCont), UserProfile(?uProf), EducationProfile(?edProf), hasConditionsProfile(?hProf, ?cProf), hasEducationProfile(?uProf, ?edProf), hasHealthProfile(?uProf, ?hProf), hasLanguage(?txtCont, ?txtLang), hasLiteracyLevel(?edProf, ?litLev), hasPatientEducation(?uProf, ?edu), hasPreferencesProfile(?uProf, ?prefProf), hasPreferredLanguage(?prefProf, ?prefLang), hasSMOGReadabilityLevel(?txtCont, ?readLev), hasSymptom(?cProf, ?sym), hasDescription(?sym, ?symDesc), hasSubject(?txtCont, ?txtSub), stringEqualIgnoreCase(?symDesc, ?txtSub), SameAs (?litLev, ?readLev), SameAs (?prefLang, ?txtLang) -> hasEducationTopic(?edu, ?txtCont)</p>
9	<p>Activity(?act), Activity(?imgAct), ActivityProfile(?aProf), HealthProfile(?hProf), ImageContent(?img), PatientEducation(?edu), PersonalProfile(?prof), UserProfile(?uProf), hasActivity(?img, ?imgAct), hasActivityProfile(?hProf, ?aProf), hasActivitySuggestion(?aProf, ?act), hasAgeGroup(?img, ?imgAge), hasAgeGroup(?prof, ?profAge), hasEthnicity(?img, ?imgEth), hasEthnicity(?prof, ?profEth), hasGender(?img, ?imgGen), hasGender(?prof, ?profGen), hasHealthProfile(?uProf, ?hProf), hasPatientEducation(?uProf, ?edu), hasPersonalProfile(?uProf, ?prof), SameAs (?act, ?imgAct), SameAs (?imgAge, ?profAge), SameAs (?imgEth, ?profEth), SameAs (?imgGen, ?profGen) -> hasEducationImage(?edu, ?img)</p>
10	<p>ConditionsProfile(?cProf), HealthProfile(?hProf), PatientEducation(?edu), UserProfile(?uProf), hasConditionsProfile(?hProf, ?cProf), hasHealthProfile(?uProf, ?hProf), hasPatientEducation(?uProf, ?edu), hasSymptom(?cProf, BlurredVision) -> hasTextSize(?edu, TxtSize_Large)</p>
11	<p>PatientEducation(?edu), PersonalProfile(?prof), UserProfile(?uProf), hasPatientEducation(?uProf, ?edu), hasPersonalProfile(?uProf, ?prof), hasEHealthLiteracyScore(?prof, ?eHealth), lessThanOrEqual(?eHealth, 20) -> hasEhealthTemplate(?edu, Template_lowHL)</p>
12	<p>Activity(?act), Activity(?uAct), ActivityProfile(?aProf), HealthProfile(?hProf), PersonalProfile(?pProf), UserProfile(?uProf), hasPhysicalActivityLevel(?aProf, Active),</p>

<pre> hasHealthProfile(?uProf, ?hProf), hasActivityProfile(?hProf, ?aProf), hasCity(?act, ?actCity), hasIntensity(?act, Moderate), hasPersonalProfile(?uProf, ?pProf), hasHomeCity(?pProf, ?pCity), hasPreferredActivity(?aProf, ?uAct), hasDescription(?act, ?actDesc), hasDescription(?uAct, ?uActDesc), equal(?actCity, ?pCity), equal(?uActDesc, ?actDesc) -> hasActivitySuggestion(?aProf, ?act) </pre>

4.6 Technical Components of the Web-Based Architecture

As illustrated in Fig. 4.1 the web-based architecture was designed as a layered architecture. Although semantic web ontology and rule languages comprised the core technologies to facilitate personalisation mechanisms, other technologies were utilised to implement the logic components, communication infrastructure and database layer. The technologies included in the architecture are as follows.

- *Java* [432] was used to develop many of the logic components, and the *OWL API* [433] Java library was used to manipulate the ontology. *JavaServer Pages (JSP)* [423] technology was used to present the education.
- The educational content was stored in a *MySQL* [434] database, and the *Spring Framework JDBC* [435] was used to retrieve content from the database.
- The architecture was hosted on an *Apache Tomcat v8.0* [436] server.

4.7 Testing

This study focused on investigating whether semantic web technologies could facilitate the creation of personalised education for diabetic patients. The architecture incorporated two semantic web components; an OWL ontology knowledge base, and a personalisation and contextualisation layer that comprised SWRL rules and a Pellet reasoner. Thereby the testing of the architecture focused on evaluating, (1) whether the OWL ontology provided an accurate and comprehensive knowledge base to support personalisation, and (2) whether the SWRL personalisation rules provided effective decision support for the personalisation. The means by which to evaluate whether these

two components could facilitate personalisation was to assess whether the architecture could produce personalised diabetic education which fulfilled the objectives described in Section 4.4.1. If the education did achieve the stated objectives, then it could be deduced that the two semantic web components did effectively facilitate the personalisation. During the testing phase various patient model profiles were developed to represent diabetic patients. These were designed to reflect a range of symptoms, treatments and complications common to diabetic patients, as well as various personal attributes, and health literacy and eHealth literacy capabilities. The design of the test profiles also aimed to ensure that all the personalisation conditions (as defined by the SWRL rules) would be evaluated. Three vignettes are presented to illustrate how personalised education can adapt to the personal and health objectives of different patients that have been diagnosed with type 2 diabetes. Accompanying each vignette is a screen shot of a section of the personalised education produced for the corresponding patient model. The vignettes were formulated to reflect the typical experiences of diabetic patients, and were based on case studies and personal experience stories as expressed by diabetic patients. The sources for these case studies and patient experience stories included a book for a diabetes healthcare team [390], booklets produced by health organisations [427, 437], and websites such as NHS Choices [66], Diabetes UK [388] and Diabetes.co.uk [438].

Vignette 1

Jane is a 65-year-old woman who has just been diagnosed with type 2 diabetes. She is worried about this diagnosis, however, her physician explains that her condition can be successfully managed through lifestyle changes and regular health checks. The surgery has a facility to provide personalised electronic education. Jane's personal, educational and health characteristics are added to her patient profile and the personalised education is generated. As Jane has been diagnosed with type 2 diabetes the personalisation rules infer that information related to blood glucose monitoring should be added to the education. Jane presented with blurred eyesight therefore the education highlights the importance of regular eye screening. Furthermore, the personalisation rules infer that the text information should be presented in an increased font size. Her educational profile is used to determine her level of health literacy and select the most suitable text components. She finds that the information has a clear flow and uses terminology that she can understand. Age and gender appropriate images are also included, which helps

Jane to identify and engage with the information. As Jane has little experience of using web-based applications the layout includes static buttons which can be used to move between the different sections of the education. Fig. 4.8 shows abridged representations of asserted and inferred individuals within Jane's patient profile, and Fig. 4.9 presents an illustration of the personalised education produced.

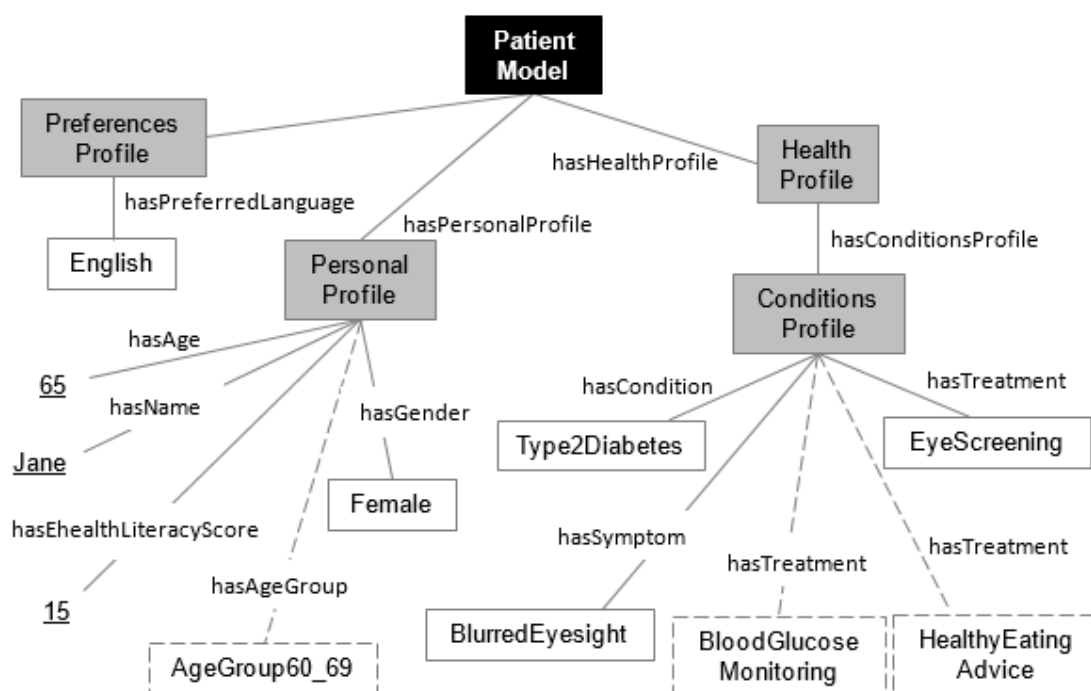


Fig. 4.8. *Patient Model instance with object and datatype properties (asserted and inferred).*

Patient has reduced eHealth literacy skills. Static link buttons are prominently displayed to aid navigation.

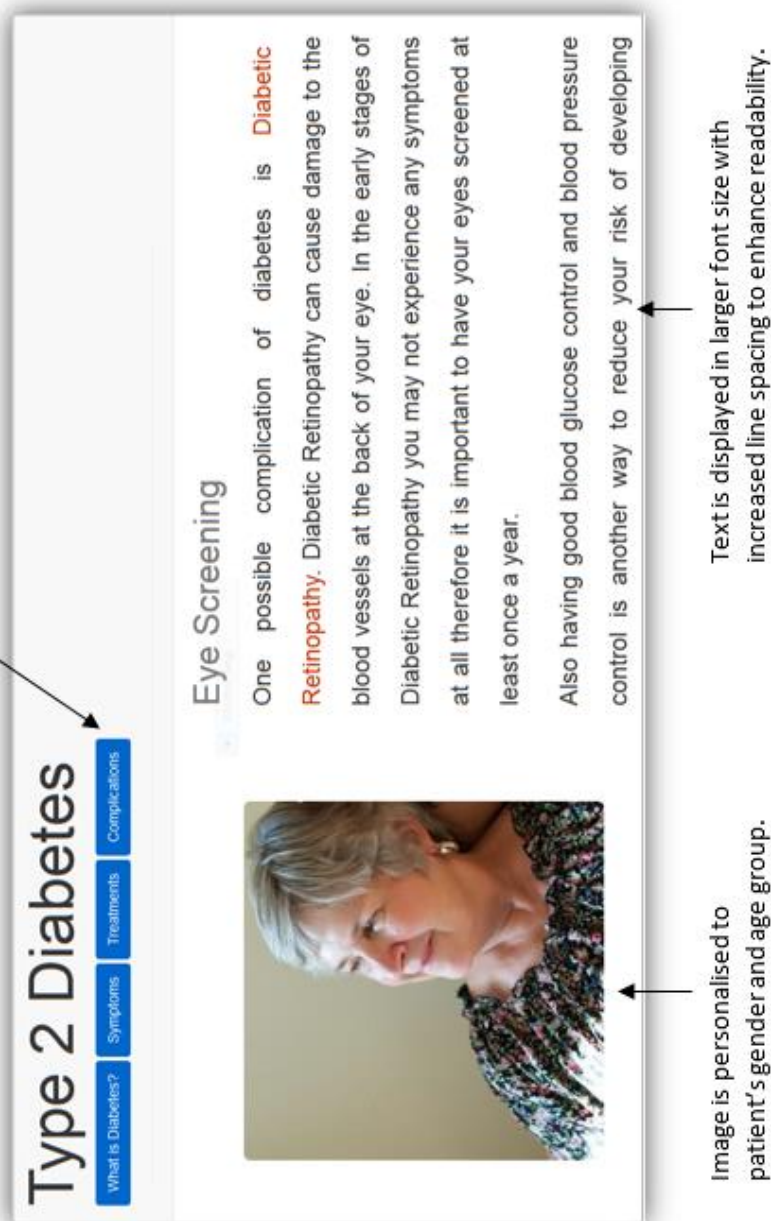


Fig. 4.9. Screen shot of personalised diabetic education for older female patient with eyesight impairment.

Vignette 2

Jack is a 31-year-old man who has just been diagnosed with type 2 diabetes. Jack is eager to increase his physical activity to help him manage his condition and has indicated cycling as a preferred activity. His personal profile holds his home address and this is used to identify cycling routes close to Jack's home. The education includes directions, a map and cycling safety tips. Jack's education also underlines the importance of regular blood glucose checks during physical activities. The images included focus on his gender and preferred sporting activities. Fig. 4.10 presents an illustration of the personalised education produced for this patient profile.

Vignette 3

Carol is a 22-year-old woman who has just been diagnosed with type 2 diabetes. Carol's BMI score is added to her patient profile and the personalisation rules infer that her BMI category is obese. Consequently, information regarding healthy eating is included in her education. Carol has decided that she will incorporate more walking into her lifestyle in order to manage her condition, however she is worried about how diabetes may affect her feet. Her physician adds foot care as a treatment in her health profile, and information related to this topic is added to her education. Carol is a university graduate and is very confident with using web-based applications. The personalisation rules infer that Carol can understand written information with a higher readability score, and may also wish to research her condition further. Thereby her education is presented in a format that is accessible for patients with advanced health literacy and eHealth literacy skills, and includes links to external accredited information resources. Fig. 4.11 presents an illustration of the personalised education produced for this patient profile.

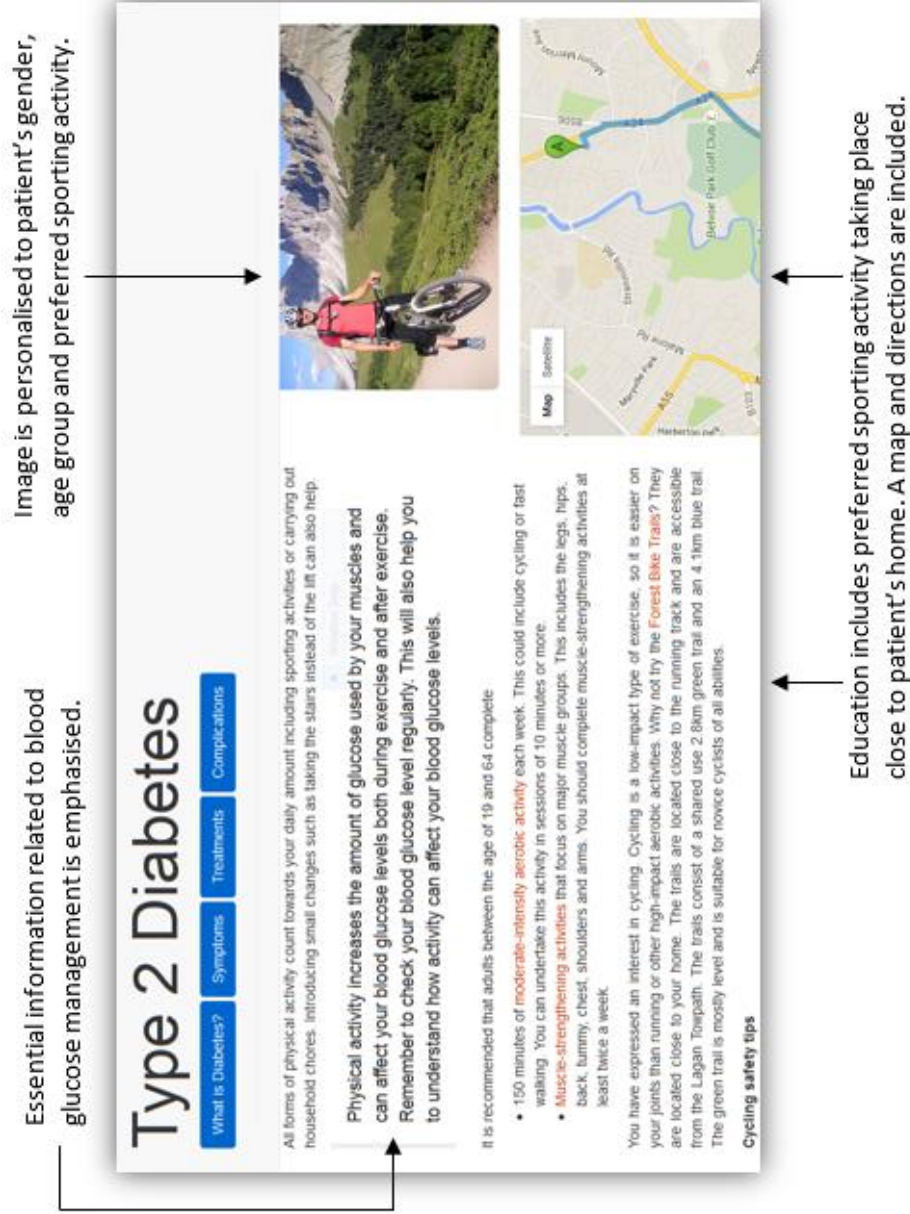


Fig. 4.10. Screen shot of personalised and contextualised diabetic education for low active male patient.

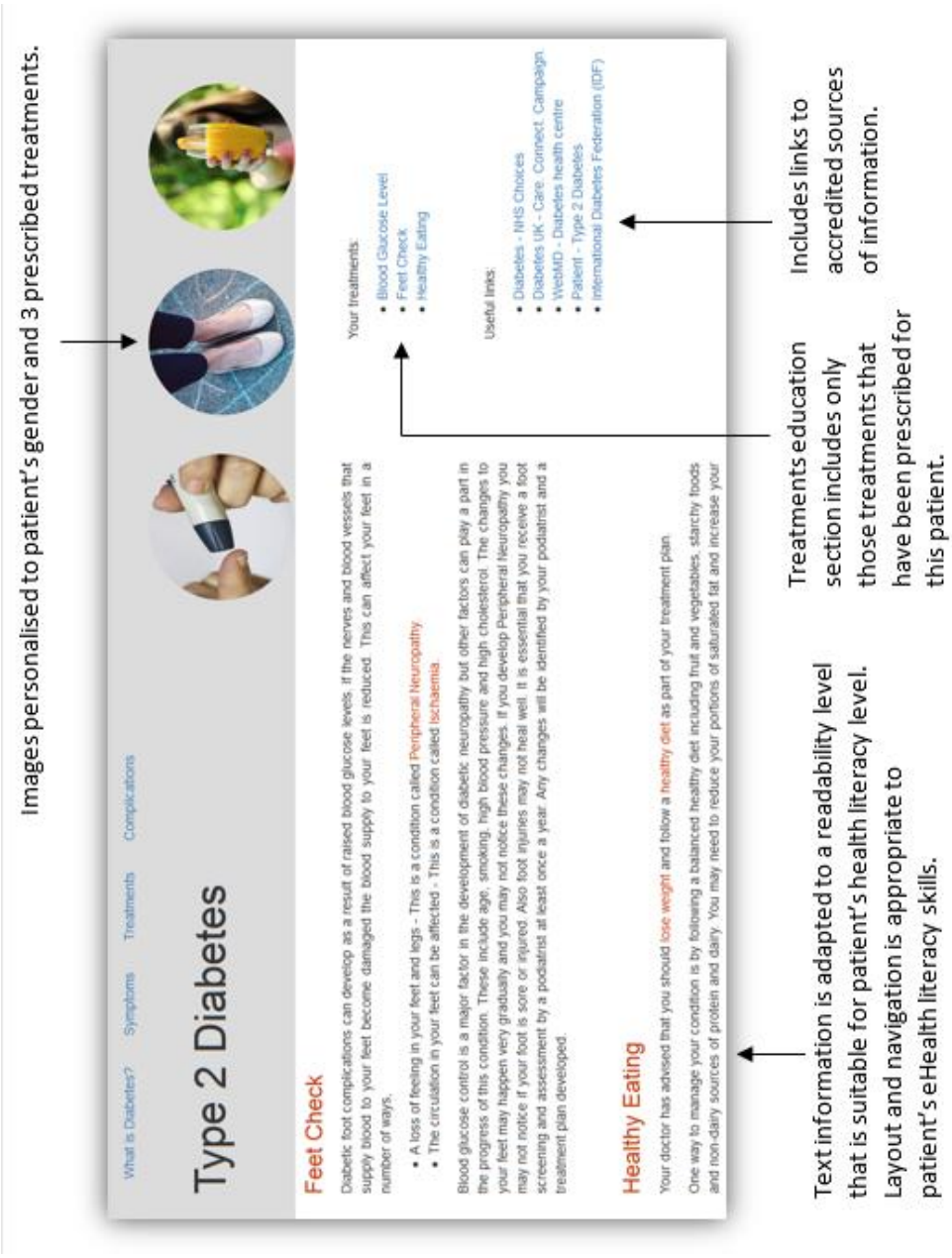


Fig. 4.11. Screen shot of personalised education for female diabetic patient with advanced health literacy and eHealth literacy competencies.

4.8 Discussion

Managing diabetes can require a patient to undertake complex daily monitoring routines, lifestyle changes and longer term health checks. Education has been recognised as an essential aspect of care for diabetes which can enhance the self-management skills of patients, and improve health outcomes [392, 395, 396]. This chapter described a study in which a web-based architecture was developed to create personalised education for diabetic patients. The aim of the study was to determine whether semantic web technologies could support the provision of personalised patient education. Central components of the architecture comprised an OWL ontology and a SWRL rule-base. Ontologies can provide the sophistication to capture the semantics of a diverse range of domain data. SWRL provides additional features to express the conditions for personalisation. Ontologies have been successfully utilised to represent a variety of clinical information related to diabetes healthcare [417-419]. Moreover, semantic reasoning is often used to produce individualised treatment recommendations. In common with these applications this study focused on developing an ontological model of clinical data related to diabetes. However, it was also required to capture patient characteristics, and information about the educational components. A review of associated literature did not identify an ontology that captured all the entities in the patient education domain, thereby it was determined to develop an original ontology for the architecture.

The domain of diabetic patient education is complex, and includes entities such as the patient, the health condition diabetes, and the actual education. Modelling domain knowledge in an ontology ensured that an accurate model of the features of each domain entity could be captured. This was achieved by creating hierarchical models which reflected the organisation of information for each domain entity. However, representing a precise model of the personalised patient education domain also required capturing relationships between the entity models. For example, both the Patient Model and the Educational Content Model have properties to define relationships with classes of the Medical Conditions Model. However, the OWL language provided a comprehensive range of features to represent intricate relationships between the entity models captured in the ontology. Using an ontology also enabled the use of semantic rules that could reason on the knowledge captured in the ontology, and draw new knowledge from the ontology. The use of SWRL enabled all the personalisation criteria for the education to

be precisely expressed, and a Pellet reasoner applied these criteria to the ontology to derive the knowledge necessary for personalisation of the education. In this manner semantic web technologies were effectively applied to support the creation of personalised diabetic patient education.

4.8.1 Issues Associated with Engineering the Patient Education Ontology

Whilst the patient education ontology functioned as expected in the architecture, it must also be acknowledged that there were several disadvantages of the chosen approach of engineering an application ontology as opposed to using a large domain ontology. Primarily, the design and development of the ontology was a time-consuming undertaking for one author, as it involved the acquisition, comparison and verification of health information from multiple sources. Possible issues may also exist with relation to the correctness of the ontology. Domain ontologies are typically verified by domain experts and ontology engineers, and validated through use in multiple applications. However, as only one author was involved in the development of the patient education ontology it is possible that content and structural errors were introduced such as misspellings, erroneous classification of domain knowledge, or omission of important domain knowledge. Moreover, scaling or maintenance of the patient education ontology could also be problematic. Health knowledge is continuously evolving and updating, and subsequently health domain ontologies require constant maintenance. However, maintenance of the patient education ontology would again be time-consuming for one author, requiring regular checks for new information and corresponding updates to the ontology. These issues may have been lessened if a large domain ontology had been incorporated into the patient education architecture instead of a locally developed application ontology.

4.9 Conclusion and Future Work

The research question queried whether semantic web technologies could support the personalisation of patient education. In addressing this question a web-based architecture was developed, which provided personalised education for diabetic patients. Personalisation of the education was facilitated by an OWL ontology knowledge base and a set of SWRL personalisation rules. The ontology captured the domain knowledge necessary for personalisation, and the semantic rules expressed the personalisation conditions. These components, combined with the use of a Pellet

reasoning engine, provided the knowledge necessary for the personalisation. This knowledge was utilised by other components of the architecture to construct and deliver personalised diabetic patient education. Thereby this study would suggest that semantic web technologies could be used to support the personalisation of patient education.

Evaluation of the completed architecture determined that web-based diabetic patient education could be adapted to suit the health needs and literacy capabilities of individual patients. Moreover, tailoring of the presentation could render the education engaging and suitable to the patient's visual preferences and eHealth literacy capabilities. One aim of personalisation is to attract and retain the attention of the user [165, 180]. The final study of this research project will investigate user engagement with the personalised patient education, in order to determine whether users are attracted to particular aspects of the personalisation. This study is described in Chapter 6.

One output from this study is an OWL ontology that represents the domain of diabetic patient education. The design of the ontology included a review of various artefacts associated with diabetes and patient education. However, evaluation is often included as a critical aspect of ontology engineering to ensure that an ontology is a comprehensive and precise model of the domain of interest. In order to ensure the accuracy of the patient education ontology a two-staged approach to evaluation was employed. The activities involved during evaluation of the ontology will be described in Chapter 5.

Chapter 5

A Two-Staged Approach to Evaluating an Ontology for Delivering Personalised Education to Diabetic Patients

5.1 Introduction

Chapter 4 described the design and implementation of a web-based architecture to provide personalised education for diabetic patients. A central component of the architecture comprised an OWL ontology which represented the domain knowledge necessary for the production of the personalised education. The ontology included a Patient Model that captured information about the patient, and a model of clinical information related to the health condition diabetes. The benefits of utilising ontologies in healthcare and biomedical systems are being increasingly recognised. An ontology provides a semantically precise and consistent model of the knowledge in a domain, in a format that is comprehensible for humans, and can be processed by reasoning technologies [256, 306, 307]. To this end ontologies are being increasingly used in healthcare and biomedical domains to express medical vocabularies, facilitate data exchange between heterogeneous systems, and provide a rigorous knowledge-base for clinical decision support systems [270, 281, 282, 284, 303, 439].

One approach to ensuring the quality of an ontological model is to include evaluation activities during the ontology development. Evaluation is an important aspect of ontology engineering which focuses on appraising an ontology, using objective criteria, in order to determine whether it reaches a quality standard [256, 300, 307]. During the design of the patient education ontology various sources of clinical information were scrutinised, so as to ensure that the ontology would represent an accurate model of the information within the diabetes domain. Furthermore, testing of the architecture indicated that the ontology provided a comprehensive knowledge base to support the creation of personalised diabetic patient education. However, in order to enhance the objectivity, preciseness and coverage of the ontology it was decided to include two phases of evaluation in the engineering methodology, as indicated in Fig. 4.2.

Ontology evaluation often includes appraisal by domain experts [440], thereby it was a prominent goal that healthcare and information technology specialists would be

involved during the first phase of evaluation. These domain experts would be engaged to complete ontology engineering tasks to define vocabulary and organisational structures for the ontology. Subsequently the knowledge contributions would be used to evaluate the accuracy of the ontology. However, it was likely that there would be variable experience of ontology engineering amongst the experts. Therefore, it was determined that a methodology would be designed that would assist ontology engineering novices to engage with a range of ontology engineering tasks. Thereby the third research question of this project is proposed.

- To what extent can ontology engineering be accessible for non-experts?

This chapter describes a study which was designed to investigate the research question. The aim of the study was to investigate whether domain experts with variable ontology engineering experience could engage with ontology engineering tasks. An essential objective of the study included the development of a methodology for engaging non-experts with a range of common ontology engineering activities. This methodology will be described in Section 5.3. Section 5.4 will describe how the knowledge contributed by the domain experts was used to evaluate and enrich the patient education ontology. During the second stage of evaluation the refined ontology was included in the web-based architecture in order to assess whether it could be used in the production of personalised patient education. Section 5.5 discusses the results of this second evaluation phase.

5.2 Related Work

As the use of ontologies in healthcare systems and biomedical research continues to grow, there is an increasing focus on establishing methods by which to evaluate the quality of an ontology [269]. Ontology evaluation, which aims to identify “*what the ontology defines correctly, does not define or even defines incorrectly*” [441], is often incorporated as an essential aspect of ontology engineering. A range of evaluation methodologies have been developed which focus on appraising quality attributes such as coverage, consistency, completeness, conciseness, expandability, and usability [281, 300, 441, 442]. Most methods of evaluation fall into one of the four classifications suggested by Brank et al. [440]; comparing the ontology to a gold standard, application-based evaluation, data-driven evaluation and appraisal by humans. Many ontology engineering projects will use multiple evaluation methods, however one of the most

common approaches involves appraisal of the ontology by domain experts using a set of established measures. Delir Haghighi et al. [443] used *criteria-based evaluation* to assess whether an ontology for medical emergency management in mass gatherings achieved its intended objectives. A focus group of 10 domain experts validated the concepts, hierarchies and relationships of the ontology and provided feedback in relation to six criteria; clarity, consistency/coherence, conciseness, expendability/extendibility, correctness and completeness. This enabled the identification of concepts that were ambiguous, inconsistent, contradictory or superfluous. In the next stage they used an automated approach to evaluating the ontology coverage. Leximancer was used to extract concepts and terms from two sets of domain documents and these were compared with the ontology. Although this proved useful for ontology validation the authors proposed that the involvement of domain expertise would also be beneficial at this stage. Bright et al. [304] implemented a two-staged approach to *intrinsic evaluation* of an ontological model of the domain of antimicrobial prescribing. They appraised the ontology classes for adherence to the design principles of Cimino's Desiderata and found that the ontology fulfilled four of the specified qualities; concept orientation, formal definitions, reject "not elsewhere classified" and context representation. In the next phase of evaluation domain experts used the laddering technique to evaluate ontology correctness. The hierarchies produced by the experts were compared to the ontology taxonomies, with high comparability between the two. Compliance with the Desiderata characteristics and the OBO Foundry's ontology design principles were also used to appraise adherence to standard ontology engineering practices for the bacterial clinical infectious diseases ontology (BCIDO) [444]. Moreover, the authors also addressed the issues associated with gold standard evaluation by devising a semi-automated approach that used clinical practice guidelines, electronic health records and expert case studies to create a collection of domain knowledge. The concepts and relationships extracted from these sources were manually compared with the BCIDO resulting in the addition of 236 terms and synonyms. The authors proposed that this approach provides an efficient means for updating a gold standard compendium and reduces the reliance on domain experts.

Application-based evaluation focuses on appraising the usefulness of the ontology as a component of a medical application [269, 300, 303]. Attributes of the application such as performance measurements or expected outputs are used to gauge whether the

ontology achieves its intended objectives. Valls et al. [407] gathered feedback from the users of an ontology based system and used this to appraise the feasibility of the underlying ontology. They developed the Actor Profile Ontology (APO) to capture knowledge related to home care assistance. This was incorporated into the K4Care Platform which was then tested by medical professionals in a real life scenario. They developed two questionnaires which were based on the Technology Acceptance Model, and used the generated scores to assess whether the ontology adequately supported certain functionalities and characteristics of the system. For example, the feedback regarding the perceived ease of use of the system was linked with system flexibility, and adaptation and personalisation capabilities, characteristics that were directly supported by the ontology. Thereby user feedback was useful in evaluating whether the ontology successfully supported particular functionalities within the system.

Ontologies are widely used within the biomedical and healthcare domains to underpin clinical decision support systems, data exchange and knowledge management. Therefore, it is essential that the ontological model is an accurate, unambiguous and consistent reflection of domain knowledge. Evaluation methods provide an opportunity to identify erroneous, inconsistent and redundant data within the ontology. Moreover, evaluation can also confirm that an ontology can achieve intended objectives within a medical system. The studies described have utilised a range of automated, semi-automated and manual evaluation procedures. In many cases human expertise was critical to confirming the validity of the ontology. In common with these approaches, domain specialists were also engaged to confirm the accuracy of the patient education ontology. However, the novelty of the evaluation methodology used in this study is derived from engaging ontology engineering novices from healthcare and information technology backgrounds to provide ontology modelling decisions. The results were collated to develop a knowledge base which was then used to validate and augment the ontology. Furthermore, application-specific evaluation of the ontology confirmed that the refined ontology achieved its functional objectives.

5.3 Methodology for Domain Expert Evaluation

As illustrated in Fig. 4.2, the evaluation of the patient education ontology was conducted in two stages. The first stage focused on constructing and using a collection of domain knowledge to evaluate the ontology. However, it was likely that not all the domain

specialists involved would have experience of using an ontology, therefore a main objective of this study was to design a methodology that would enable ontology engineering novices to contribute their domain knowledge. A set of ontology engineering tasks were developed which collected knowledge related to ontology classes, properties and restrictions, and were presented to participants through an online survey. This approach to ontology evaluation had three main objectives.

- (1) To evaluate the content and organisation of the patient education ontology.
- (2) To enrich the patient education ontology with new knowledge.
- (3) To determine the feasibility and efficacy of using ontology engineering novices to evaluate a health ontology.

5.3.1 Design of an Ontology Modelling Screencast

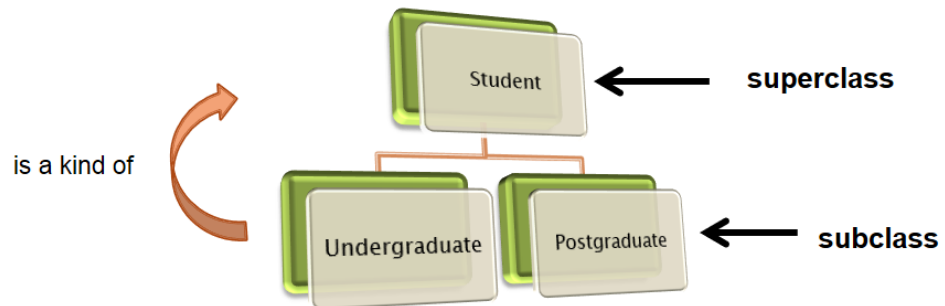
It was expected that there would be variation in the technical modelling expertise of the participants, therefore a challenge lay in developing tasks that would render ontology engineering accessible to non-experts. Consequently, an online multimedia presentation was developed that would provide the participants with the necessary information. The presentation used illustrations and analogies to cover technical topics such as organising domain information as classes, using subsumption hierarchies to encapsulate specialisation, and using relationships to model associations between unrelated classes. The presentation also introduced the diagram style and terminology that would be used in the survey. Fig. 5.1 illustrates three excerpts from the presentation. In Fig. 5.1 (a) an analogy is drawn with family relationships in order to clearly explicate ontology superclass and subclass relationships, whilst Fig. 5.1 (b) highlights that subsumption relationships can have more than one level. These screens also introduced the terminology, “*is a kind of*”, which was used throughout the survey to assist with defining superclass and subclass relationships. Fig. 5.1 (c) discusses how relationships can exist between unrelated classes, and introduces the concept of object properties. A multimedia presentation was chosen as the medium for conveying this information as it enabled a visual presentation and voiceover that would be more engaging in comparison with textual information. The screencast time was limited to three minutes so that a participant would not be overloaded with unnecessary information.

(a)

Developing a class hierarchy

Superclass-subclass relationship

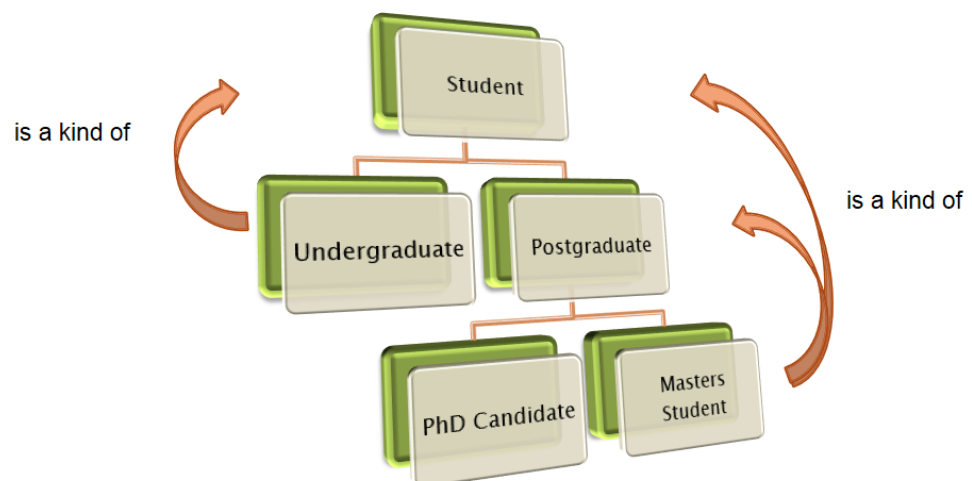
- ▶ This type of relationship is similar to a parent child relationship.



(b)

Superclass-subclass relationships

- ▶ Superclass-subclass hierarchies can have more than 1 level.



(c)

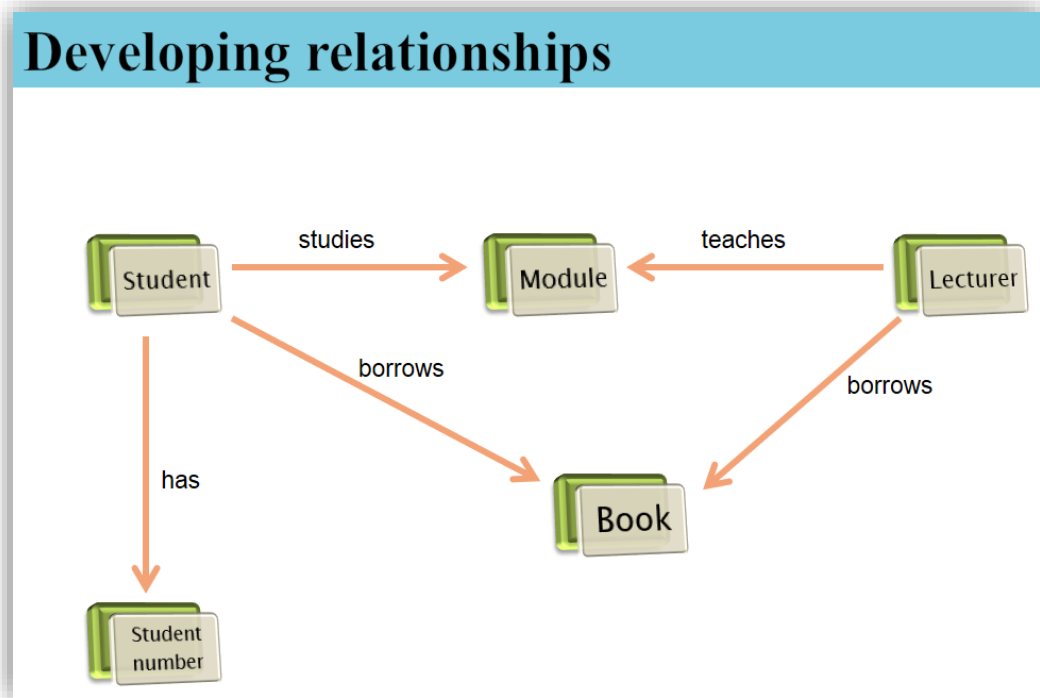


Fig. 5.1. Excerpts from online multimedia presentation focusing on (a) superclass-subclass relationships, (b) subsumption relationships with multiple layers, and (c) relationships between unrelated ontology classes.

5.3.2 Design of the Ontology Engineering Tasks

The survey focused on collecting knowledge related to three of the main domain entities which included, (1) the patient, (2) the health condition diabetes, and (3) physical activities. Obesity is a common comorbidity of diabetes with various overlapping clinical features, therefore it was decided that the survey would also collect information about the symptoms, treatments and complications of this health condition. In order to reduce the complexity of the survey it was organised into sections that focused on the following subjects.

- Patient health characteristics.
- Physical activities.
- General information related to diabetes and obesity.
- Symptoms.
- Treatments.

- Complications.

The ontology engineering tasks were designed to reflect five types of ontology modelling practices.

- (1) Suggest new data to be captured in the ontology and assign a cardinality.
- (2) Enter terms and organise these in a subsumption hierarchy.
- (3) Enter terms and classify these within an ontological model.
- (4) Analyse relationships in the current ontology and suggest alternative representation.
- (5) Enter new properties for a class.

The design of the survey was tailored to assist with the modelling tasks. Fig. 5.2 illustrates how layout and terminology were used to assist with transcribing technical modelling decisions. For example, Fig. 5.2 (a) exemplifies a set of tasks in which a participant is required to model subsumption relationships; the participant can enter a subclass name and a superclass name in the corresponding text boxes to arrange in a subsumption relationship. The terminology used, “*is a kind of*”, was introduced in the screencast as illustrated in Fig. 5.1 (a) and (b). Fig. 5.2 (b) illustrates a set of tasks in which the participant is required to enter new terms and classify these in the ontology. As illustrated, the participant can enter a symptom definition and select an option to indicate how it should be organised in an ontological model. Fig. 5.2 (c) illustrates a task in which the participant is required to describe a patient health characteristic, and choose an option to indicate the cardinality of this piece of information in the ontology.

The survey also contained descriptions, examples and ontological models to assist with modelling decisions. Health information was also included to contextualise the information being requested. For example, definitions of physical activity intensities were provided to assist with a task (illustrated in Fig. 5.2 (d)) in which the participant categorised a list of sporting activities in an ontological model. This information was gathered from various sources including the UK National Health Service [66] and Diabetes UK [388].

(a)

7. Can you suggest any subclasses of class Diabetes Medication?

Subclass

is a kind of Superclass

1.

is a kind of

(b)

2. Can you suggest any symptoms and indicate how these should be classified in the ontology?

Obesity

Diabetes

Both

1.

Symptom

(c)

1. Can you suggest any health characteristics that should be captured for a patient? How many pieces of information would be captured for each patient?

Exactly 1

At least 1

1.

Health Characteristic

(d)

1. For each activity indicate the class to which it should belong.

Light intensity

Moderate intensity

Vigorous intensity

Aerobics

Basketball

Canoeing

Fig. 5.2. Screen shots of ontology engineering tasks that focus on (a) subsumption relationships, (b) classifying terms within an ontological model, (c) assigning cardinality, and (d) classification of classes within an ontological model.

5.3.3 Profile of Invited Contributors

The online tasks required knowledge of diabetes and obesity, and the ability to identify associations between the entities in these domains. Consequently, it was decided to invite participants from two main disciplines, Information Technology and Health Sciences, as participants from these backgrounds would provide a diversity in the necessary expertise and skills required for the tasks.

5.4 Results of Domain Expert Evaluation

The survey comprised 15 questions that collected demographic information, 18 ontology engineering tasks and a comments section. The full set of ontology engineering tasks is available in Appendix D. The demographic information collected included age, gender, and background discipline. In addition, the participants indicated whether they had previous experience of using ontologies, databases, and programming, scripting or mark-up languages.

5.4.1 Participants

Demographic information was completed by 26 participants but only 21 undertook the ontology engineering tasks. The criteria for inclusion in the study required that a participant completed at least one of the ontology based tasks, therefore five participants were excluded from the study. The remaining subjects comprised 15 (71.4%) male and six (28.6%) female participants, with an age range of 24 to 70, and a mean age of 42.48 years (SD 12.7). The background disciplines of the participants included Information Technology (76.2%, $n = 16$), Mathematics and Engineering (14.3%, $n = 3$), and Healthcare, and Life and Health Sciences (9.5%, $n = 2$). Only seven participants (33.3%) had previously used or developed an ontology, six of whom had an Information Technology background and one a background in Mathematics. The main ontology languages used by these participants included OWL and RDF, while Protégé was the only ontology editor used. The majority of participants (90.5%, $n = 19$) had experience of using a database, this including one participant with a background in Healthcare. A similar proportion of participants (85.7%, $n = 18$) had experience of using mark-up languages, this including one participant from a Life and Health Sciences discipline. Seventeen of the participants (80.9%) had experience of using programming or scripting languages, all of whom were from an Information Technology, Engineering or Mathematics background.

Fig. 5.3 indicates the participants' self-ratings of expertise in information technology, and Fig. 5.4 the expertise ratings for diabetes and obesity. There was a high level of information technology expertise, however this was to be expected as 90.5% ($n = 19$) of the participants had a technical or mathematics related background, and all these participants rated their expertise as 7 or above. The two remaining participants from health-related disciplines rated their expertise as 6 or below. In contrast, all of the expertise ratings for diabetes and obesity were scored as 7 or below, with the exception of one participant (from an Information Technology background) who rated their diabetes expertise as 9.

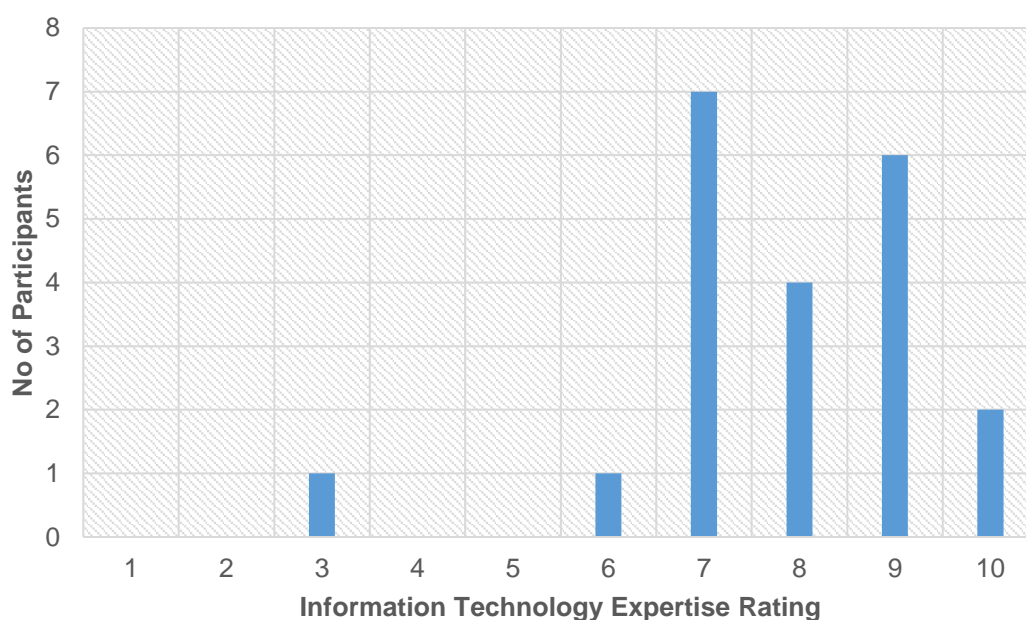


Fig. 5.3. *Distribution of information technology expertise amongst study participants.*

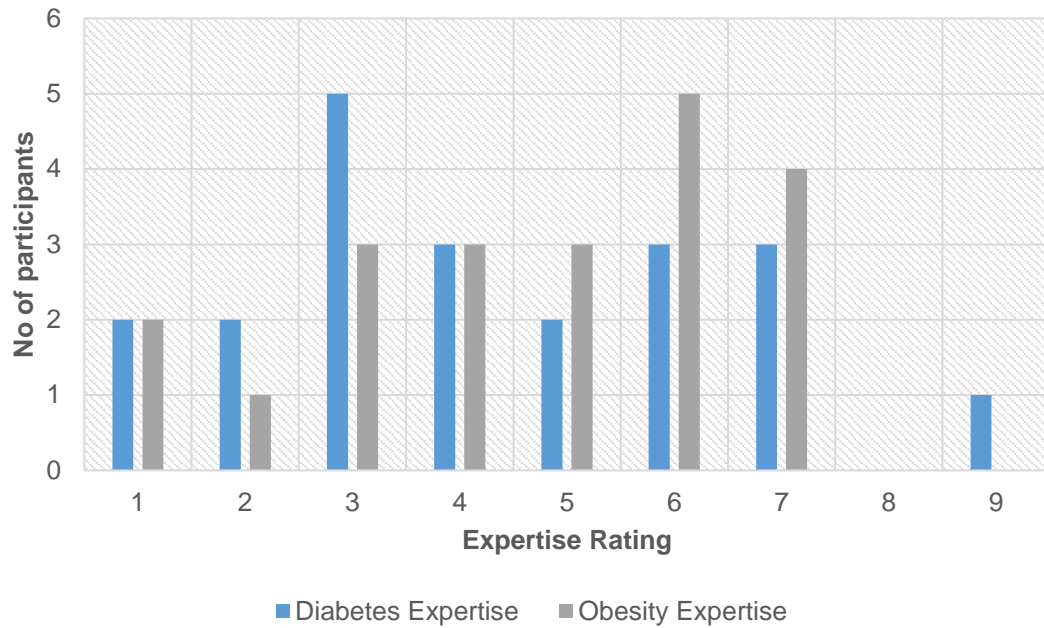


Fig. 5.4. *Distribution of expertise for diabetes (blue) and obesity (grey) amongst study participants.*

A primary objective of this study was to appraise the feasibility of engaging ontology engineering novices with ontology engineering activities. The viability of engaging the domain experts with ontology engineering tasks was appraised through the following observations.

- Analysis of the knowledge provided by the participants through the ontology engineering tasks in order to determine whether the contributions could be utilised to evaluate or enrich the patient education ontology.
- Analysis of the completion rate of the different types of technical modelling tasks by the participants.

5.4.2 Analysis of the Knowledge Contributions from Ontology Engineering Tasks

In total 936 knowledge contributions were collected from the ontology engineering tasks. The contributions included terms to define classes in subsumption relationships, new object and datatype properties, and analysis of existing hierarchical structures in the ontology. Fig. 5.5 provides a visualisation of the most frequently contributed terms throughout the survey.

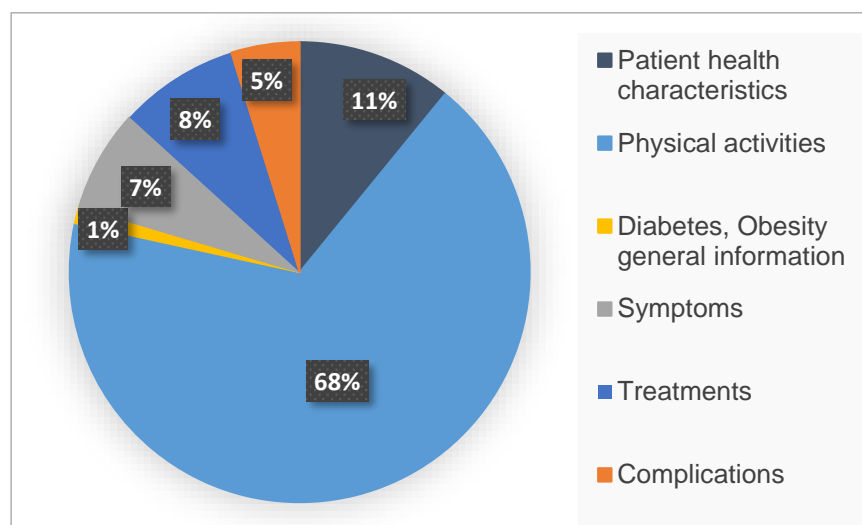


Fig. 5.6. Pie chart indicating percentage of total knowledge contributions for survey sections.

Fig. 5.6 indicates the percentage amount of the total knowledge contributions for each of the survey sections. The set of tasks related to *Physical activities* collected the largest portion of contributions (68%, $n = 632$), and this was largely due to the inclusion of a task in which the participants categorised sporting activities by intensity (as illustrated in Fig. 5.4 (d)). This was the most frequently completed task of the survey (completed by 20 participants) and collected 559 (59.7%) contributions. The second most frequently attempted task related to *Patient health characteristics* (as illustrated in Fig. 5.4 (c)). This task was attempted by 18 (85.7%) of the participants and collected 102 (11%) contributions. The section which collected general information about diabetes and obesity had least contributions. However, this section contained only one task (which was attempted by 11 participants) which may have contributed to the low rate of knowledge collected. The section related to *Complications* collected less contributions (5%) when compared to the sections that focused on *Symptoms* (7%) and

Treatments (8%). In total 12 participants (57.1%) did not attempt any of the tasks in this section. However, the *Complications* section was the final section of the survey, thereby it may be possible that participant fatigue or diminishing engagement may have contributed to the lower contribution rate.

5.4.2.1 Symptoms, Treatments and Complications

Fourteen of the tasks collected information related to the symptoms, treatments and complications of diabetes and obesity. During processing the contributions were aggregated and similar suggestions were identified and clustered. Following this, the ontology was searched for a corresponding class or property. If a match was found, then this was interpreted as validation of the existing ontology. If a match was not found a literature search was conducted in order to determine if the suggestion could be verified as a feature of diabetes or obesity. The sources reviewed included the UK National Health Service [66], Diabetes UK [388], MedlinePlus [68] and websites that had been accredited with a Health on the Net certification [162]. If the contribution was verified as an aspect of diabetes or obesity it was added to the ontology, otherwise it was categorised as invalid.

Table 5.1 illustrates the accumulated knowledge contributions for each of the survey sections and how these were used to validate and augment the ontology. In total 67 (7.2%) of the contributions related to the symptoms of diabetes and obesity, and these included generic conditions and specific indications of diabetic complications. The most frequently suggested symptom was fatigue which was classified as a symptom of both diabetes and obesity, as too was high BMI, and shortness of breath. The most recurrent symptoms specific to diabetes included blurred eyesight, frequent urination, excessive thirst, unexplained weight loss, dry mouth, unbalanced blood sugar level, abdominal pain, disorientation and fainting. All these contributions were matched with symptoms represented in the ontology. A number of new suggestions were also added as classes in the ontology. This included heart racing as a symptom specific to hypoglycaemia, and high cholesterol as a symptom of obesity. Most significantly there was a high level of overlapping between the *Symptoms* and *Complications* sections as approximately 15% of the suggested symptoms were verified in the literature as diabetes complications. This would suggest that some participants were unclear of the distinction between diabetes symptoms and complications.

Table 5.1. Accumulated knowledge contributions by subject, with brackets indicating percentage of total contributions.

Topic	Total Contributions	Mapped to ontology	Added to ontology
Patient health characteristics	102 (11%)	84 (8.9%)	4 (0.4%)
Physical activities	632 (67.5%)	605 (64.6%)	14 (1.5%)
Diabetes, Obesity general information	11 (1.2%)	6 (0.6%)	1 (0.1%)
Symptoms	67 (7.2%)	52 (5.6%)	6 (0.6%)
Treatments	79 (8.4%)	51 (5.4%)	7 (0.7%)
Complications	45 (4.8%)	28 (3.0%)	11 (1.2%)
Totals	936	826 (88.2%)	43 (4.5%)

The majority of the contributions for diabetes and obesity treatments were focused on non-medication treatments. The most commonly suggested generic treatments included healthy diet and increased exercise. Diabetic health checks and monitoring routines were also recommended including daily blood glucose measurement, HbA1c blood test, blood pressure measurement, eye screening, foot checks, weight checks and recommended annual health checks. All these suggestions were mapped to classes in the patient education ontology. A number of suggestions related to mental health assistance. A literature search confirmed that this should be added to the ontology as a treatment. Approximately 4% of the contributions related to medication treatment. Insulin was the most common suggestion and it was also proposed that information relating to insulin delivery methods and regimes should be represented in the ontology. These were added as properties of the Insulin classes. Metformin was the only non-insulin medication mentioned which suggests that there could be a lack of knowledge related to non-insulin medications amongst the participants. In summary, 5.4% of the contributions were used to validate the patient education ontology and 0.7% were used to add new knowledge.

The contributions related to the complications of diabetes and obesity included both physical and psychological problems. The suggestions included breathlessness, depression, blurred vision, foot ulcers, cardiovascular disease, joint and back pain, poor

circulation and amputation. Moreover, while some of the contributions used general descriptions of problems, a number of complications were expressed using precise medical terms including neuropathy, retinopathy and sleep apnoea. All of these contributions were mapped to classes in the patient education ontology. Blurred vision had been represented in the ontology as a symptom of diabetes, however a literature search verified that it should also be classified as a complication. A number of new complications were proposed such as loss of mobility, foot pain and heart attack. These were added as classes in the ontology. There were also suggestions that recording contextual information or behaviour that increased the risk of a complication should be recorded. This was represented in the ontology as properties of patient behaviour in the patient profile. Overall 3% of the contributions were used to validate the current ontology and 1.2% were utilised to add new classes and properties.

5.4.2.2 Physical Activities

One of the main approaches to managing diabetes and obesity is through increased physical activity, however it is important that any sporting information provided to a patient is appropriate to their physical activity level. The first task in the *Physical activities* section required the participants to assign physical activity intensities to a list of 28 sports. Majority voting was used to determine the most appropriate intensity from the contributions. In most cases there was agreement amongst the participants, however there were conflicting assignments for three of the sports. In these cases a reference sporting compendium [445] was utilised to determine the most appropriate intensity. Fig. 5.7 illustrates how the classifications were used to update the model of physical activities, and also includes 12 new sporting activities which were proposed by the participants. Eight activities of daily living (ADLs) were also suggested which included various gardening and housework activities. These were also added as new classes in the ontology.

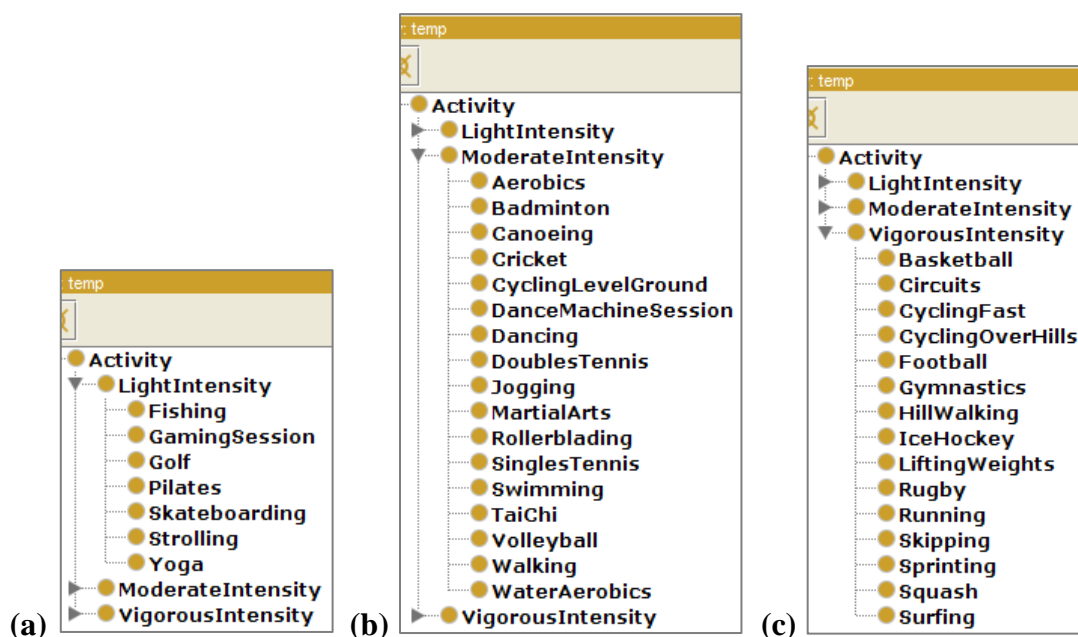


Fig. 5.7. Subsumption relationships within the enhanced physical activity model, including classes to represent (a) light intensity activities, (b) moderate intensity activities, and (c) vigorous intensity activities.

5.4.2.3 Patient Health Characteristics

In order to gather information that could be used to enhance the patient profile, the participants were asked to suggest patient characteristics that could be relevant to personalised education. Moreover, the participants also had to indicate the cardinality of this characteristic. Table 5.2 indicates the ten most frequently suggested patient health characteristics. All of these, and further suggestions including ethnicity, gender, preferred language and literacy level, were mapped to classes and properties in the ontology. Novel suggestions comprised information related to the patient's medical team. This was added to the ontology as it could be useful for inclusion in personalised patient education.

Table 5.2. Ten most frequently suggested patient health characteristics and associated cardinalities, with brackets indicating rank order.

Patient health characteristic	Cardinality		Total number of unique suggestions
	Exactly 1	At least 1	
Exercise Information		✓	10 (1)
Weight	✓		10 (1)
Blood Pressure Reading	✓		9 (2)
Comorbidities		✓	9 (2)
Blood Sugar (Glucose) Level	✓		7 (3)
Cholesterol Level	✓		6 (4)
Diet Information		✓	6 (4)
Age	✓		5 (5)
BMI Score	✓		5 (5)
Height	✓		4 (6)

5.4.2.4 Summary

As denoted, the knowledge contributions enabled the evaluation of various models of domain knowledge within the patient education ontology. In summary, 88.2% of the knowledge contributions were used to validate the content and organisation of the ontology, and 4.5% were used to add new classes and properties. The total quantity of classes in the ontology increased by 16% and the total object properties increased by 18%.

5.4.3 Participant Engagement with the Ontology Engineering Tasks

This study aimed to evaluate whether ontology engineering novices could make valuable contributions to ontology evaluation. Section 5.4.2 described how the knowledge provided by the participants through the ontology engineering tasks was used to evaluate and expand the coverage of the patient education ontology. Analysis of the knowledge contributions would suggest that ontology engineering non-experts can make valuable knowledge contributions to ontology evaluation processes. The next

stage of analysis focused on the participants' level of engagement with the different types of ontology engineering tasks in the survey. The average survey completion time was 31 minutes which would suggest that the level of engagement amongst the participants was high. However, in order to further investigate whether the participants were willing to engage with ontology modelling activities, the completion rates for the different types of ontology engineering tasks were evaluated.

5.4.3.1 Completion of the Technical Tasks

The success of the methodology for ontology evaluation was hugely reliant on the participants' motivation to engage with the ontology engineering tasks. As described in Section 5.3.2 the survey comprised five types of tasks that focused on general ontology modelling activities. Therefore, in order to appraise the participants' willingness to undertake technical tasks the completion rate for each task type was calculated. Table 5.3 indicates the number of participants from each discipline that attempted each task type.

The first task type required a participant to assign a cardinality (exactly 1, at least 1) to a data property. As indicated in Table 5.3, this question was completed by 90.5% ($n = 19$) of the participants, including both participants from a health-related discipline. This would suggest that participants with a range of technical expertise would be confident with this data modelling practice.

The second task type used subsumption relationships to organise data. Approximately 42.9% ($n = 9$) of the participants did not attempt any of these tasks. Moreover, only three of the seven participants with experience of using an ontology attempted this task type. This would suggest that conceptualising and organising classes in hierarchical relationships may be challenging for ontology engineering novices, including those with comparable technical knowledge.

The third task type focused on providing and classifying information; the participant was shown a hierarchical model from the ontology and asked to classify information in different classes. Only 1 participant (from a health-related background) did not attempt any questions of this type. Thereby, such a high completion rate (95.2%, $n = 20$

participants) would suggest that participants from both technical and non-technical backgrounds may be able to classify information within ontology classes.

Table 5.3. Completion rate of ontology engineering tasks, with brackets indicating percentage of participants from discipline group.

Task No.	Ontology engineering task type	No. of participants that completed task type			
		Information Technology	Mathematics & Engineering	Healthcare & Life Sciences	Total
1	Enter data, assign cardinality	14 (87.5%)	3 (100%)	2 (100%)	19 (90.5%)
2	Enter terms in subsumption relationship	9 (56.3%)	2 (66%)	1 (50%)	12 (57.1%)
3	Enter terms and classify	16 (100%)	3 (100%)	1 (50%)	20 (95.2%)
4	Analyse current relationships	11 (68.8%)	1 (33%)	1 (50%)	13 (61.9%)
5	Enter class properties	7 (43.8%)	1 (33%)	1 (50%)	9 (42.9%)

The fourth task type required the participant to analyse an ontological modelling decision. Approximately 61.9% (n = 13) of the participants attempted these tasks. In parallel with the second question type (subsumption relationships), this would suggest that some participants may find it difficult to assess the suitability of hierarchical structures for representing information in an ontology.

The fifth task type focused on adding new properties for a class. This had the lowest completion rate with only 42.9% (n = 9) of the participants attempting these tasks. Moreover, less than half of the participants with a technical background attempted any of these tasks. It was surprising that this task type had such a low completion rate as it

was similar to the first task type (enter data, assign cardinality) which had a high completion rate (90.5%, $n = 19$ participants).

Overall there was a disparity amongst the participants in willingness to attempt the different task types. In total eight participants (38.1%) attempted all five task types. This included one participant with a healthcare background, and one from an engineering background, neither of whom had used an ontology before. The remaining six participants had an information technology background, two of these had previous experience of using an ontology and all had expertise in using databases and scripting languages. It is notable that a participant from a healthcare discipline attempted all of the ontology engineering task types. This would suggest that ontology evaluation need not be limited only to domain experts with technical expertise, but that valuable technical contributions can be collected from non-technical personnel.

Some of the more unexpected findings of the study related to the participants that had previous experience of using or developing an ontology. Over half of these participants did not attempt any type 2 (subsumption relationships) or type 5 (suggest new properties) tasks, and only 57% ($n = 4$) attempted any type 4 (analyse current relationships) tasks. As indicated in Section 5.4.1 the majority of participants had indicated previous experience of using high level programming languages and databases, therefore it was decided to explore whether there were any associations between particular technical experience and willingness to undertake the different types of ontology modelling tasks. Fig. 5.8 illustrates the percentage of participants from each of the technical knowledge groups (defined by previous technical experience) that attempted each task type. In accord with the previous analysis (which focused on background discipline), task types 2 and 5 appeared the most problematic as less than 50% of participants from each of the technical knowledge groups attempted these. However bivariate analysis of the data found no statistically significant relationships between previous technical experience and willingness to undertake the different types of ontology engineering task.

Analysis of the knowledge contributions also examined whether there were any relationships between the participants' contribution totals and the expertise ratings for information technology, diabetes or obesity. Fig. 5.9 indicates the average number of

knowledge contributions collected from each expertise category. Surprisingly some of the lowest averaged contributions were from participants that had rated their diabetes expertise in the upper quadrant. However, this may be associated with a reluctance from some participants to engage with the ontology modelling tasks as three of the four participants with the highest rated diabetes expertise (7-9) did not attempt any of the type 2 (subsumption relationships) or type 4 (analyse current relationships) ontology engineering tasks. However, again bivariate analysis of the data did not identify any statistically significant relationships between the expertise ratings and level of knowledge contributions.

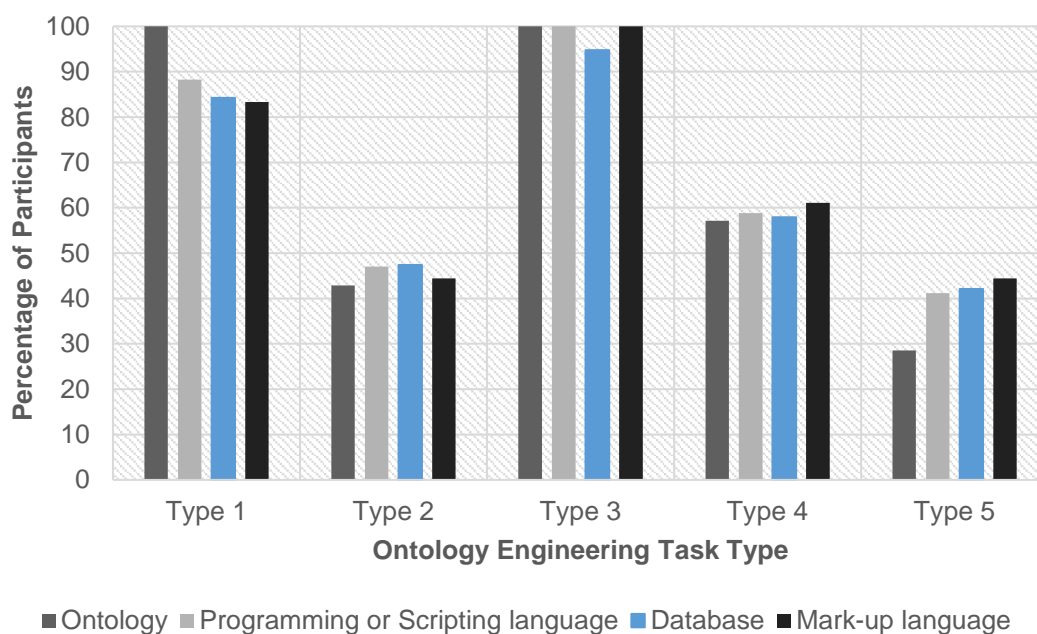


Fig. 5.8. *Percentage of participants that completed tasks, organised by sub-groups defined by previous technical experience of using an ontology (dark grey), a programming or scripting language (light grey), a database (blue), and a mark-up language (black).*

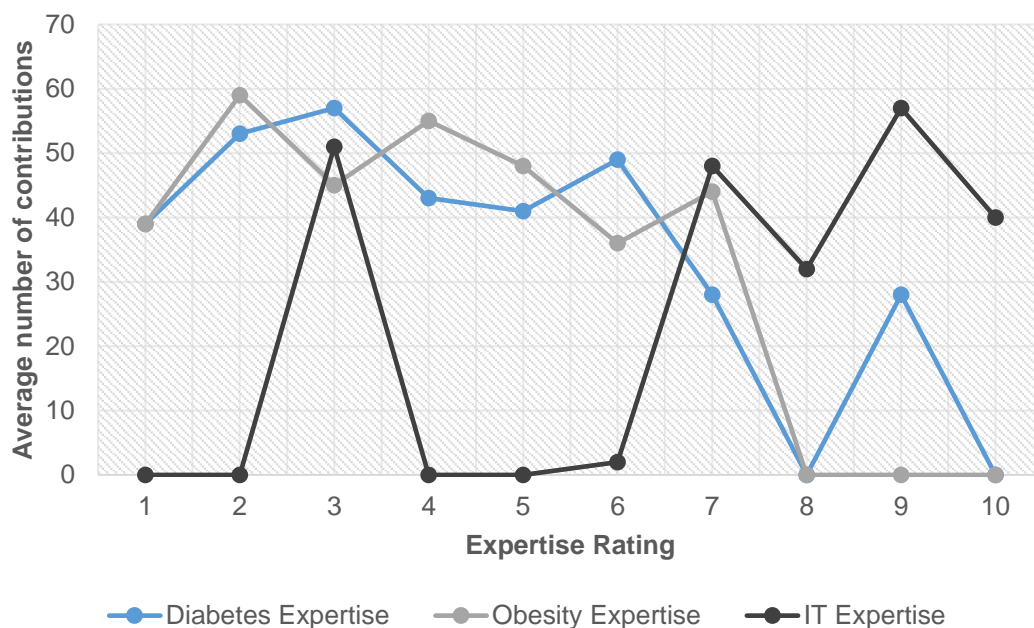


Fig. 5.9. Averaged contributions related to expertise rating categories for diabetes (light blue), obesity (grey), and information technology (black).

5.5 Application-Based Evaluation

The second stage of evaluation aimed to assess whether the enhanced ontology could be used in the web-based architecture to facilitate the production of personalised patient education. A test suite of Patient Model instances were developed to model patients that had been diagnosed with type 2 diabetes. The test profiles included common symptoms, treatments and complications of type 2 diabetes, and were designed to reflect a diversity of gender, age and educational abilities amongst the patients. Section 5.5.1 presents a vignette related to one of the test profiles. The following criteria were used to evaluate whether the personalised education corresponded with a test profile.

- The education includes the patient's name.
- The text components of the education describe only the particular symptoms, treatments and complications as asserted in the patient's profile or inferred by the personalisation rules.
- The text is at an appropriate readability level as specified by the patient's educational profile.
- The images chosen are matched to the patient's gender and age.
- The education layout is as specified in the patient's preferences profile.

5.5.1 Test Profile

John is a 35-year-old man who has just been diagnosed with type 2 diabetes. He is worried about this diagnosis, and as his BMI category is overweight, he is particularly keen to manage his weight as a means to help manage his condition. Consequently, his physician prescribes increased physical activity and healthy eating as part of his treatment. John's surgery has a facility to provide personalised electronic education, and his physician explains that information personalised to his specific treatment plan is accessible through a web-based portal as John requires. John's personal, education and health characteristics are added to his patient profile along with his treatment plan. As John has been diagnosed with type 2 diabetes the personalisation rules infer that information related to blood glucose monitoring should be added to the education. Moreover, John's BMI score is utilised to infer that information related to healthy diet should be added to his education also. His physical activity profile specifies golf as a preferred sport therefore his home address (stored in his personal profile) is used to identify and provide information about golf courses close to his home. From John's education profile it is inferred that he has a high level of health literacy, and as he has indicated a high level of eHealth literacy (in his preferences profile) the readability and layout of his education is adapted accordingly. Fig. 5.10 presents an illustration of the personalised education produced for this patient profile.

5.5.2 Results of Application-Based Evaluation

In all of the test cases the education matched the characteristics as specified in the profile, and all the evaluation criteria were fulfilled. Each patient had been diagnosed with type 2 diabetes therefore the personalisation rules inferred that information related to blood glucose measurement should be added. In each case the education focused only on each patient's particular experience of symptoms, treatments and complications. The text information was provided at a suitable readability level and the images corresponded with the patient's gender and age. The results of the application-based evaluation indicated that the enhanced ontology provided a comprehensive knowledge base to assist with the production of personalised patient education.

Images personalised to patient's gender and treatments.

What is Diabetes? Symptoms Treatments Complications

Type 2 Diabetes

Blood Glucose Level

Most treatments for diabetes help to keep your blood glucose levels as normal as possible. This means that you should try to keep your blood glucose levels in the range of a person who does not have diabetes. This should help to control your symptoms and prevent health problems or diabetic complications developing. You may also hear blood glucose levels referred to as **blood sugar levels**.

If your blood glucose level is too high this may lead to a condition called **hyperglycaemia**. If your blood glucose level is too low this may lead to a condition called **hypoglycaemia**.

If you do not monitor your blood glucose level this may damage your body and can lead to diabetic complications.

To test your blood glucose at home prick the side of one finger with a finger-pricking device. Then put a drop of blood on a testing strip. This will give you an accurate reading of your current blood glucose level. Your diabetes health team will work with you to find the most suitable testing device.

Your treatments

- Blood Glucose Level
- Healthy Eating
- Increased Activity

Useful links

- Diabetes - NHS Choices
- Diabetes UK - Care, Connect, Campaign
- WestMD - Diabetes health centre
- Patient - Type 2 Diabetes
- International Diabetes Federation (IDF)

Readability of text information is suitable for patient's health literacy level, and layout is appropriate to patient's eHealth literacy skills. Educational content focuses on treatments that have been asserted (increased activity) and inferred (blood glucose monitoring, healthy eating) for this patient.

Fig. 5.10. Screen shot of personalised diabetic education for male patient which includes information focused on asserted and inferred treatments.

5.6 Discussion

Evaluation is an essential phase of ontology engineering through which the accuracy and correctness of an ontological model can be enhanced [256, 300, 440]. There are many methodologies available for ontology evaluation [281, 300] and most projects will use more than one evaluation method to ensure the validity of the ontology. Medical literature and clinical documents are often scrutinised in order to derive evidence based domain knowledge from which an ontological model of a domain can be constructed or evaluated [418, 421, 444, 446]. A similar methodology was employed during the development of the patient education ontology as a review of printed and online medical literature was undertaken in order to develop a knowledge base related to diabetes. Domain experts are often employed during evaluation processes to appraise the correctness and conciseness of an ontology, or to develop a reference gold-standard to be used for evaluation. As domain experts were not involved in the design of the patient education ontology, a prominent aim for the evaluation phase was to use domain expertise to appraise the accuracy and correctness of the ontology. Healthcare and information technology specialists were engaged to complete ontology engineering tasks from which a collection of domain knowledge was gathered. This knowledge was then utilised to evaluate the ontology. Expert led ontology evaluation often includes the development of relationship models of an ontology, for example by use of the laddering technique [304, 444]. In a similar manner, some of the participant contributions were utilised to define subsumption relationships which were used to confirm the accuracy of the ontological relationships within the patient education ontology. The majority of contributions matched the constructs within the ontology, and were used to verify the accuracy of the ontology. This outcome is comparable to domain expert evaluations for other medical ontologies in which the experts were largely in agreement with the ontological models developed for the application [304, 444]. However, the quantity of novel suggestions collected from the participants was very low as only 4.5% of the contributions comprised new knowledge which was added to the ontology. Consequently, there was little change to the content and structure of the ontology through the evaluation process.

The main novelty in this approach to ontology evaluation is derived from engaging domain experts that had variable comprehension of ontologies. Artefacts that described various attributes of ontologies, and intuitive interfaces that assisted with engineering

tasks were provided. The majority of participants attempted a range of modelling activities, this including participants that had no previous knowledge of ontologies, and in particular one participant from a health background whom had never used a database or programming language. This would suggest that some domain specialists with limited modelling experience can successfully complete technical modelling tasks and can make valuable contributions to ontology evaluation. However, examination of the completion rate for the different types of ontology engineering tasks suggests that the participants may have found some tasks especially challenging. In particular, a notable quantity (42.9%) of the participants did not attempt any of the tasks related to modelling subsumption relationships. This included participants that had previous technical experience of using ontologies, databases and high level programming languages. This would suggest that modelling information in subsumption relationships may prove a complex undertaking for ontology engineering novices, even for those with a technical background. The lowest completion rate was for type 5 tasks which required the participant to suggest new properties for various ontology classes. A similar task had been completed by a high percentage of the participants (90.5%) which would suggest that the majority of participants had an understanding of class properties. However, it may be possible that the low completion rate may be associated with the participants' knowledge of diabetes and obesity. Most of the type 5 tasks focused on suggesting properties (information) related to the symptoms, treatments and complications of diabetes and obesity, thereby it is feasible that a lack of knowledge may have impacted the level of participant contributions in these tasks. Participant comments suggested analogous reasons for performance during the survey; one participant commented that lack of time prevented them from providing more contributions, whilst four others indicated that a lack of knowledge of diabetes and obesity had the same effect. However, as the study comprised an anonymous online survey the motivations of some participants to avoid certain types of ontology engineering task cannot be clearly ascertained. Bivariate analysis of the data did not uncover any statistically significant relationships between willingness to undertake certain types of tasks and participant characteristics such as background discipline or technical knowledge.

The high degree of overlap between the contributions and the ontology would suggest that the study was successful in validating the accuracy of the patient education ontology. However, it must also be acknowledged that this high level of agreement may

be related to the characteristics of the study population. The majority of participants ($n = 19$, 90%) had a technical background and thereby may have had a similar technical profile to the ontology author. This may have impacted the results of the study in several ways. Firstly, it is likely that many of these participants had technical modelling knowledge that was comparable to that of the ontology author. Consequently, some of these participants may have agreed with the modelling decisions made by the author (such as the hierarchical ontological models) thereby contributing a modelling decision that matched the organisation of the current ontology. Secondly, awareness of personalisation techniques may also have contributed to an overlap between the participants' contributions and the ontology content. For example, the second most frequently attempted task related to the collection of patient health characteristics that would be used to personalise the education. It is likely that subjects with a technical background would have been aware of the characteristics that are typically included in user profiles to support personalisation, and thereby would have suggested characteristics that matched those included by the ontology author.

There are a number of benefits that can be derived from this methodology for engaging novices with ontology engineering. Firstly, the scope for engaging domain expertise is expanded, as involvement in evaluation activities is not limited to those with ontological modelling experience. Furthermore, as the contributions were collected through an online survey this was a time-efficient approach to collecting domain expertise. However, the collation and validation of the contributions was manually conducted by the investigator. Future work could focus on automation of these processes, thereby increasing the efficiency of this evaluation methodology.

The second phase of evaluation focused on the usefulness of the enhanced ontology within the web-based architecture. The results of this evaluation indicated that the ontology provided an accurate knowledge base for the production of personalised diabetic patient education.

5.6.1 Study Limitations

There were a number of limitations with this study. The number of participants that engaged with the online survey was relatively small, and most notably participation from the healthcare domains was very limited. If a greater number of participants from

a healthcare background had engaged with the survey, then the range and specificity of the diabetes clinical information collected may have been enhanced. This may have enabled evaluation of a greater proportion of the patient education ontology, or may have further augmented the coverage of the domain.

5.7 Conclusion

The research question focused on the extent to which ontology engineering could be accessible for individuals with variable experience of using or building an ontology. In order to answer this question a methodology was designed to assist ontology engineering novices with completing ontology engineering tasks. The knowledge contributions collected through the tasks were utilised to validate the accuracy of the patient education ontology and to expand the coverage of the domain. The survey collected a total of 936 contributions, 88.2% of which were used to validate the ontology and 4.5% were used to add new knowledge to the ontology. This resulted in a 16% increase in the number of classes in the ontology and an 18% increase in the number of object properties. However, due to the high level of similarity between the diabetes knowledge contributions and the ontology, the ontological models of symptoms, complications and treatments remained largely unchanged. Further analysis of the knowledge contributions also highlighted variation amongst the participants in willingness to attempt different types of task. Whilst a number of participants with a technical background engaged with a range of ontology engineering tasks, only one participant from a health background undertook all five types of modelling tasks. Moreover, completion rates highlighted that certain types of ontology modelling practises may have been challenging for some participants.

Several challenges were encountered during this study. There was a very low level of participation by individuals from health-related disciplines, which likely contributed to the lack of variation and specificity of diabetes clinical knowledge amongst the contributions. Moreover, this reduced the potential to analyse the ontology modelling capabilities of individuals with non-technical backgrounds. Most of the subjects were from technical disciplines and it is possible that the high level of agreement between the contributions and the ontology content could be related to the technical knowledge of the Computer Science participants.

The outputs from this study would suggest that there is variation in the capability of ontology engineering novices to contribute to ontology engineering processes. Although some of the participants did not engage with all the modelling tasks, analysis of the completion rate would suggest that ontology engineering can be accessible for non-experts if sufficient information is provided and if an individual is motivated to engage with modelling tasks. However, in some sections of the survey there was a low rate of contributions and a lack of variety in the diabetes clinical data suggestions. Moreover, in other sections there was a high level of overlapping in the contributions. Consequently, the amount of new knowledge added to the ontology was limited. Moreover, as many of the contributions matched the content of the ontology, there was little change to the ontological models of diabetes symptoms, complications and treatments. In conclusion, the results of this study would suggest that ontology engineering practices may be accessible for some novices that have sufficient knowledge of a health domain and are interested in technical modelling. However, ontology engineering may also prove challenging both for individuals with technical modelling experience and for persons from a non-technical background.

5.7.1 Issues Associated with Engineering Health Ontologies

Chapter 4 and Chapter 5 described the development of the patient education ontology. This section highlights several issues associated with engineering health ontologies, and proposes some recommendations.

This chapter described an approach to evaluating the patient education ontology, however attempts to appraise the ontology were limited by a lack of diabetes clinical knowledge amongst the study participants. This was disappointing as it was hoped that the study would provide a thorough evaluation of the correctness and accuracy of the ontology. Health knowledge is characterised by complexity, inconsistencies and variability, and it can be difficult to conceptualise this knowledge in an ontology [447, 448]. For instance, the use of synonymous terminology is common in many healthcare domains. As an example, many diabetes medications have both a brand name and a generic name. Representing synonyms can add complexity to an ontology, if two or more entries have the same meaning. Achieving a precise representation of complex clinical information can also be challenging. For instance, OWL DL object properties represent binary relationships whereby an assertion can link only two individuals.

However, some aspects of health knowledge can necessitate higher degree associations to be represented [298]. For example, the condition of an anatomical organ may change as a disease progresses, requiring that a temporal dimension or inter-relationships with other concepts be recorded [447]. These examples highlight some of the issues associated with recording health information in an ontology, and underline the value of having access to specialist domain knowledge during engineering processes. Thereby the first recommendation is to involve health domain experts during the design, development and evaluation phases of ontology engineering. As highlighted, health knowledge can be intricate and multifaceted, and therefore professional assistance is beneficial to ensure that the health knowledge has been modelled correctly and that all corresponding properties and relationships have been accurately represented. If possible, several domain experts should be engaged in the engineering processes as this will reduce the possibility of bias.

The second recommendation is related to the design of ontology-based applications. The patient education ontology provided an accurate knowledge base for the personalised education architecture, however, developing the ontology was time-consuming and it is probable that similar models of diabetes were available in several large domain ontologies. Moreover, a domain ontology could also have been extended if more specific class and relationship definitions were required for the personalised education application. Thereby the second recommendation is that when designing an ontology-based application, the ontology authors should consider the use of an established domain ontology instead of developing an application specific ontology. Several online portals are available which provide access to a large collection of health and biomedical ontologies which cover a wide range of health domains. Employing a domain ontology may lessen the time required to develop an ontology-based application, by reducing the time and labour required to design, develop and evaluate a suitable ontology.

5.8 Future Work

Chapters 4 and 5 described the design, implementation and evaluation of various components of a web-based architecture that produces personalised diabetic patient education. It has been suggested that personalised information can be more engaging for an individual when compared with generic information [167]. Thereby the final

study of this research project focuses on examining user engagement with the personalised patient education, and aims to investigate whether users are attracted to particular components of the education. This study is described in Chapter 6.

Chapter 6

An Evaluation of User Engagement with Generic and Personalised Patient Education for Diabetic Patients

6.1 Introduction

Patient education is an essential component of healthcare services as it can equip patients with the knowledge and skills they require to manage their condition and make informed decisions about subsequent treatment and daily care [449]. However, whilst access to printed and online health education has proliferated, many health consumers experience reduced ability to engage with the information acquired [450]. As highlighted in Chapters 2 and 3, health literacy is a major determinant of a patient's skills to understand, retain and apply health information. However, while health literacy may impact an individual's capability to use health information, the quality attributes of health education artefacts may also reduce their comprehensibility for health consumers [65]. For many patients printed educational materials such as leaflets and booklets provide a trusted form of health information [21, 24, 65]. Printed materials can provide standardised, medically certified information that can be revisited at varying stages of treatment [81, 449]. However, as highlighted in Chapter 2, various studies have found that the content, format and readability of written artefacts do not always match the information needs and literacy competencies of all patients [65, 78, 84-87]. There are concerns that if educational materials are incongruent with patients' literacy skills then some patients may misunderstand the information given or disengage with further healthcare [114].

In order for a patient to process and subsequently recall health information they must be attentive to the health information provided [58, 165]. However, a challenge exists in identifying methods of providing health information that will encourage patients to engage with the material. Various reviews of health behaviour change interventions have described how tailored health communications have been effective in motivating health consumers to engage with positive behavioural changes [169, 170, 451]. One means by which tailoring could be applied to attract the attention of a patient is to present health education as self-referential. Personalisation is a technique that incorporates recognisable characteristics of an individual (e.g. an individual's name)

into a health message in order to enhance the relevance of the information for that person [165, 173]. Perceiving a message as self-relevant may intensify central processing of the message by an individual [165, 170, 173]. Chapter 4 described the development of a web-based framework that provides personalised education for diabetic patients. Various personalised features were included in the education so as to assist a patient to identify and engage with the education. However, while testing of the framework indicated that the personalisation mechanisms produced accurately tailored education, it was unclear whether the personalised education would be successful in attracting the attention of a patient. Thereby the fourth research question of this project is proposed.

- Do any relationships exist between the format of education being used and the engagement behaviour exhibited by an individual?

This chapter describes a study which aimed to investigate the research question. The aim of the study was to explore the engagement behaviour exhibited by an individual when using either a generic education booklet or personalised diabetic patient education. The first study objective was to identifying whether there were any similarities or contrasts in the engagement patterns of individuals when using the different types of health education. During the study, groups of participants read either a generic booklet or personalised education as produced by the patient education framework (as described in Chapter 4). Viewing behaviour was recorded and utilised, along with the participants' self-measures of engagement, to uncover patterns of engagement with the two formats of education. The second study objective addressed participant engagement with the web-based personalised education, and examined whether there were any differences in the level of attention directed towards the different personalised features of the education. Section 6.3 describes the methodology used in the study. This includes a description of the two different educational artefacts, and how eye tracking was used to measure engagement with the education. The results of the study are reported in Section 6.4 and further discussed in Section 6.5. Section 6.2 discusses related research in which eye tracking has been used to examine how health consumers read and utilise health information.

6.2 Related Work

Having access to health information can assist an individual to become more knowledgeable about their health needs and better able to exercise informed choices about their healthcare [9, 27, 450]. However, for a health consumer to be able to process health information they must find it comprehensible [450]. Tailoring has been proposed as an effective strategy for adapting health communications to the varied literacy competencies and information needs of patients. For example, Attack and Luke [452] describe a patient education website which enables a physician to tailor the information provided, by choosing education artefacts, such as text and video, which are directly applicable to a patient's information needs. Users of the website indicated that the tailored information was useful and helped them to improve their knowledge. The authors of [453] investigated whether applying different levels of personalisation to educational materials could assist acquisition of knowledge. They found that materials that were adapted to the health literacy needs and learning preferences of patients were associated with the greatest improvements in participants' knowledge.

Chapter 4 described the development of a framework to provide personalised education for diabetic patients. The content, readability and presentation of the education was tailored to the personal characteristics, preferences, health literacy skills and health objectives of a patient. This chapter describes the next stage of evaluation which focused on engagement with the education. *User Engagement* is a multifaceted concept for which various definitions, activities and behaviours have been proposed [454]. For example it has been suggested that user engagement is revealed through a user's interest and attention towards technology [454]. O'Brien and Toms [454, 455] proposed that engagement "*is a quality of user experiences with technology*" which is characterised by various dimensions including aesthetic and sensory appeal, interest, motivation and usability. They developed the *User Engagement Scale (UES)* [456], a self-report instrument by which users could indicate their engagement with a system with relation to attributes such as aesthetic appeal, felt involvement, perceived usability, focused attention, endurability and novelty. The UES has been validated as a reliable measure of engagement in various online platforms including news and ecommerce [456, 457].

A contrasting approach to evaluating an individual's attention towards health information is through analysis of eye gaze activity. Eye tracking is a technique for

documenting and quantifying eye movements [458], and has been successfully employed to appraise users' focus of attention and cognitive load when searching for and using health information [58, 459-461]. Eye movement data can indicate where the user's interest is directed on a visual display and how the level of interest fluctuates in response to different visual stimuli. Different types of eye movements are evaluated to identify the pattern of user gaze behaviour. The most prominent types of movement used include, (1) a *fixation*, during which the eye is stationary and focused on a specific area of a visual display, and (2) a *saccade*, in which the eye moves rapidly between fixations [458]. An individual will perceive visual images only during periods of fixations, and are wholly unable to see during saccades [458]. It has been proposed that the length of a fixation is indicative of the level of cognitive processing being applied towards an object [458, 462].

Health risk information is often displayed on the packaging of consumer goods such as alcohol and cigarettes. Consequently, much research has focused on examining whether this information is noticeable, and is read by consumers. For example, Kessels and Ruiter [459] describe an eye tracking experiment that evaluated the level of attention that smokers and non-smoker directed towards health risk information and health coping information on cigarette packets. Analysis of fixations and dwell time indicated that smokers attended more to coping information regardless of the accompanying images. In contrast non-smokers focused on coping information when it was accompanied by a low-risk image, but spent more time observing risk information when a high-risk image was shown. An eye tracking investigation of attention towards brand information and health information on alcohol and soda packaging found that the subjects spent a larger amount of time viewing branding information [460]. Nevertheless, for all the subjects the size and appearance of alcohol labels did have an impact on viewing time as alcohol warning labels that were larger and less complex were attended to for longer periods.

Eye tracking methodologies have also been utilised to investigate the associations between health consumer characteristics and the manner in which health information is used and recalled. Mackert et al. [461] used eye tracking to investigate whether there was an association between the information viewed on the NVS nutritional label and health literacy competency. They found that subjects with reduced health literacy skills

(as denoted by lower NVS scores) tended to fixate on non-relevant information for longer periods of time when compared to subjects that had more advanced health literacy capabilities. A study which examined attention towards a text only webpage and a text-illustrated webpage found no association between health literacy competency and the fixation rate or fixation length for the two formats of health information [58]. However, it was found that increased attention to the text of the webpage correlated with increased recall of the information for those with adequate health literacy, but that an equivalent relationship was not found for those with limited health literacy. Conversely increased attention to the illustrations was associated with improved information recall for those with limited health literacy skills but not for participants that had adequate health literacy skills. The authors of [463] used eye tracking to examine the gaze behaviour of low literacy adults when viewing either a factsheet produced by the Centers for Disease Control and Prevention (CDC), or a literacy appropriate “dirty bomb” decision aid which included images and less text. They found that the more densely worded factsheet was more challenging for some participants. Fixation analysis indicated that some participants tended to dwell on certain words and phrases for longer durations, whilst gaze patterns showed that some participants had an unsystematic approach to reading the text, repeatedly re-reading lines of text or looking at the text in a haphazard manner.

Health education comprises a process which may assist patients to become more activated in decisions about their health. However, the education provided may not be fully accessible for all patients. The reviewed literature suggests that health consumers will employ different methods to engage with and process health information. Moreover, some individuals will not be attentive to all the information provided. Thereby strategies are required that will motivate patients to attend to health education. This chapter describes a study which aimed to appraise engagement with generic and personalised formats of diabetic patient education. User engagement is driven by various characteristics thereby a dual approach to measuring engagement was employed. The UES provided a means to gather subjective ratings of engagement from the participants themselves, whilst eye tracking afforded evaluation of the participants’ visual responses to different aspects of the education. This enabled evaluation of the participants’ experience of the visual appeal and usability of the education, whilst also appraising which features of the education attracted most attention.

6.3 Methodology

During the study each participant attended an eye tracking experiment. The protocol followed during an experiment is indicated in Fig. 6.1. The eye movements of the participants were tracked during stages 2, 3 and 4. In the first stage a participant completed demographic information including age, gender and highest qualification, and rated their knowledge of diabetes on a ten-point Likert scale of one to ten (ranging from very limited to expert). The participant also rated their information technology expertise on a ten-point Likert scale of one to ten (ranging from novice to expert). Following this the eye tracking began, and the participant read diabetic patient education which was presented on a computer screen. The type of education displayed (generic or personalised) was dependent on whether the participant had been assigned to the control or experimental group. No time limit was set to complete reading the education. When reading was complete the participant was presented with four health questions, and was free to use the education to find answers for the questions. Following this the participant completed the UES.

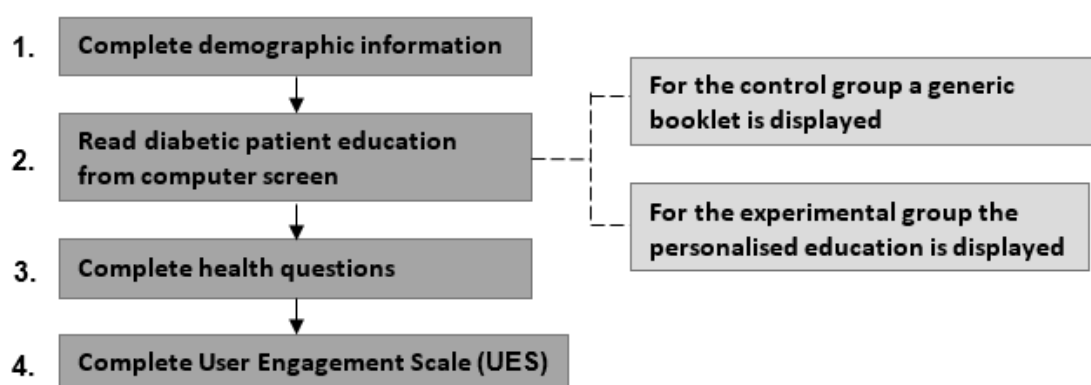


Fig. 6.1. Protocol for eye tracking experiment.

6.3.1 Participants

A total of 18 subjects participated in the study, which comprised 10 (56%) male and eight (44%) female participants. The background disciplines of the participants included Computer Science (50%, $n = 9$), Politics (11.1%, $n = 2$), Museum Studies (5.6%, $n = 1$), Arts (11.1%, $n = 2$), English (5.6%, $n = 1$), and Cookery (5.6%, $n = 1$). Two participants did not list a discipline.

This was a between subjects study in which a *control group* of nine participants viewed a *generic education booklet* for diabetic patients, and an *experimental group* of nine participants viewed *web-based personalised diabetic education*. Before the experiments began each participant was randomly assigned to one of these groups. The control group included six (66.7%) male and three (33.3%) female participants, with an age range of 24 to 36, and a mean age of 31.1 years ($SD = 4.31$). All of the members of this group had achieved an undergraduate or postgraduate degree. The experimental group included four (44.4%) male and five (55.6%) female participants, with an age range of 19 to 49, and a mean age of 32.9 years ($SD = 9.61$). Six (66.7%) of the participants in this group had attained an undergraduate or postgraduate degree, and the remaining three participants (33.3%) had a Regulated Qualifications Framework qualification at level 2 or level 3 [361]. Fig. 6.2 indicates the participants' self-ratings of expertise in information technology, and Fig. 6.3 the level of diabetes knowledge. As can be perceived, the range of information technology and diabetes expertise within the two groups was similar. The average diabetes knowledge score for the control group (3.33 out of 10) was greater than the average score for the experimental group (2.89 out of 10), however the difference between the two scores was not statistically significant ($t_{(16)} = -.476, p = .641$). Similarly, the mean score for information technology expertise was greater for the control group (6.56 out of 10) when compared to the experimental group (5.89 out of 10), however again the scores did not differ significantly ($t_{(16)} = -.514, p = .616$).

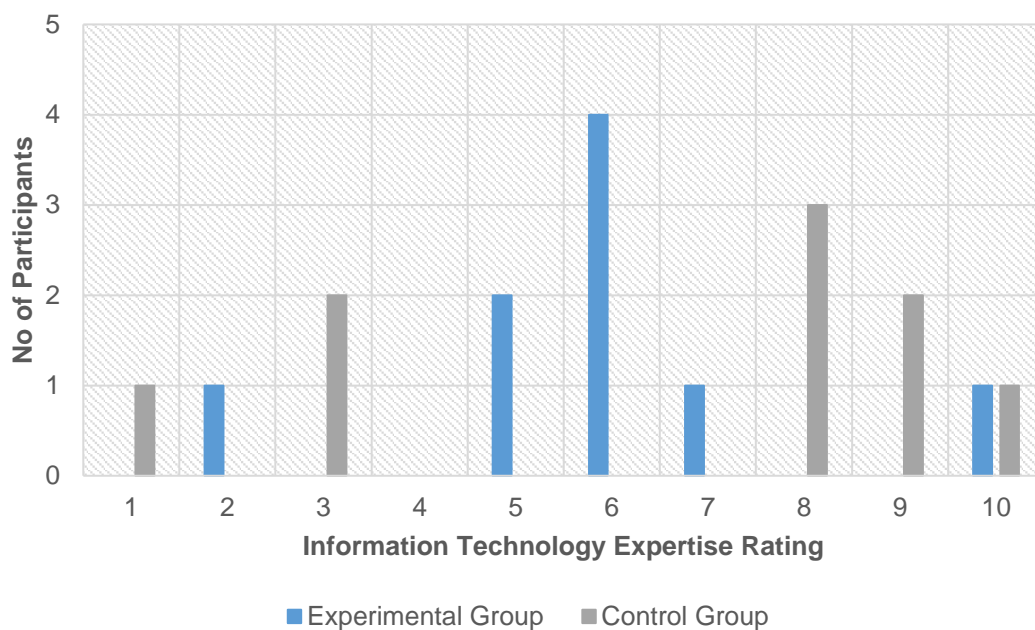


Fig. 6.2. Information technology expertise ratings for control and experimental groups.

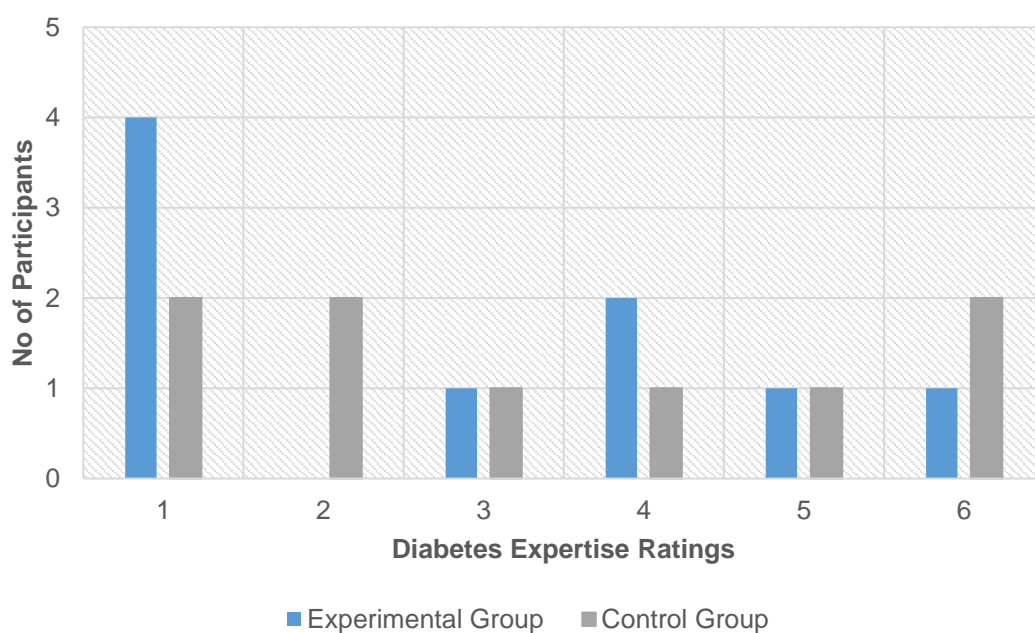


Fig. 6.3. Diabetes expertise ratings for control and experimental groups.

6.3.2 Diabetic Patient Education

A main objective of this study was to investigate engagement patterns with generic and personalised patient education. During the study the participants that were assigned to the control group viewed a generic education booklet, and members of the experimental

group viewed web-based personalised education. The main characteristics of these two education artefacts are now described.

6.3.2.1 Generic Patient Education Booklet

Members of the control group viewed an introductory booklet for diabetic adults, *Understanding diabetes. Your essential guide* [464], which was published by *Diabetes UK*, a registered UK charity. This booklet is aimed at patients that have been diagnosed with type 1 or type 2 diabetes, or caregivers of diabetic patients. The participants were required to read nine pages of the booklet which covered the following topics.

- What is diabetes?
 - Type 1 and type 2 diabetes.
- Diagnosis.
 - Coming to term with diagnosis.
 - Lifestyle changes and choices.
- Managing diabetes.
 - Blood glucose levels, blood pressure and blood fats, medication.
 - Eating well.
 - Keeping active.
- Short-term complications.
 - Hypoglycaemia.

The booklet contained text and images. For example, Fig. 6.4 illustrates a page of the booklet that focuses on *keeping active*. The topics addressed were clearly identified by headings and sub-headings. The background of the booklet was white, thereby the written content (black text) and headings (pink text) were clearly visible. The nine pages read by the participants contained in total nine images. The content of these pages did not alter during the study thereby all the members of the control group were presented with the same text and images.

The participants were required to read nine pages of the booklet. The average number of words on a page was 225.22 words. The Simple Measure of Gobbledegook (SMOG) readability formula [429] was used to determine the readability of the text content. The

SMOG score was 9.8, and this readability level was understandable for all the members of the control group (whom were all university graduates).

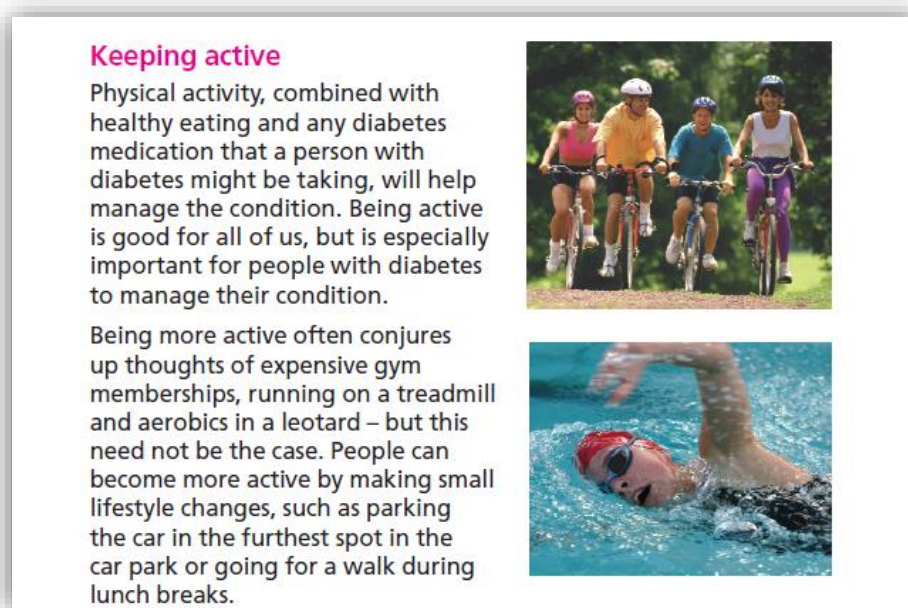


Fig. 6.4. Portion of a page from the generic education booklet [464] which focuses on physical activity.

6.3.2.2 Web-Based Personalised Patient Education

Members of the experimental group viewed web-based personalised education. This education was produced by the personalised patient education framework which was described in Chapter 4. The education consisted of five web pages which focused on a sub-section of the topics that were included in the generic booklet. Fig. 6.5 indicates the structure of the personalised education and the subject matter focus of each of the five web pages. Links were included on each web page so that participants could navigate back and forth between the pages in whichever order they preferred.

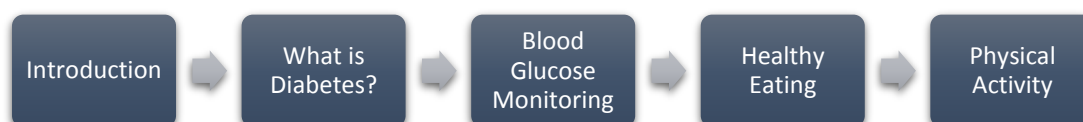


Fig. 6.5. Structure of web-based personalised education indicating subject matter of the five web pages.

The personalised education incorporated both *static* features which remained constant for all the participants, and *personalised* features, which varied in accordance with the characteristics of a participant. For example, the layout of the web pages and the text, headers and sub-headers remained largely consistent for all the participants. Four of the web pages (2 - 5) also included at least one *static image* which remained unchanged for all the participants. The personalised features of the education were as follows.

- The first line of the first web page (introduction) comprised a personalised greeting which included the participant's name. This feature was included in order to enhance the self-relevance of the education for a participant.
- Each of the five web pages contained at least one *personalised image* which was automatically selected based on the participant's personal characteristics and preferences. These were included so as to encourage the participant to identify with the education. For example, Fig. 6.6 illustrates a screen shot of the *physical activity* web page which contains information for a female participant. The two personalised images are gender and age specific and are also contextualised to the preferred sporting activity of the participant.
- The *physical activity* web page contained information that had been personalised to the preferred sporting activity of the participant, and included a Google map which highlighted a route from the participant's current geolocation to a location for the sporting activity.

The subject matter of the web pages was signposted by use of headings and sub-headings, and readability was enhanced through the use of paragraphs, bullet points, white-space and columns. Written text was coloured black and thereby clearly visible on a background of either white or light blue. In summary, the personalised education contained six static images, seven personalised images (which includes the personalised map). The average number of words per web page was 356.6 words. The SMOG readability score for the education was 7.6, thereby the text content was comprehensible for all the members of the experimental group.

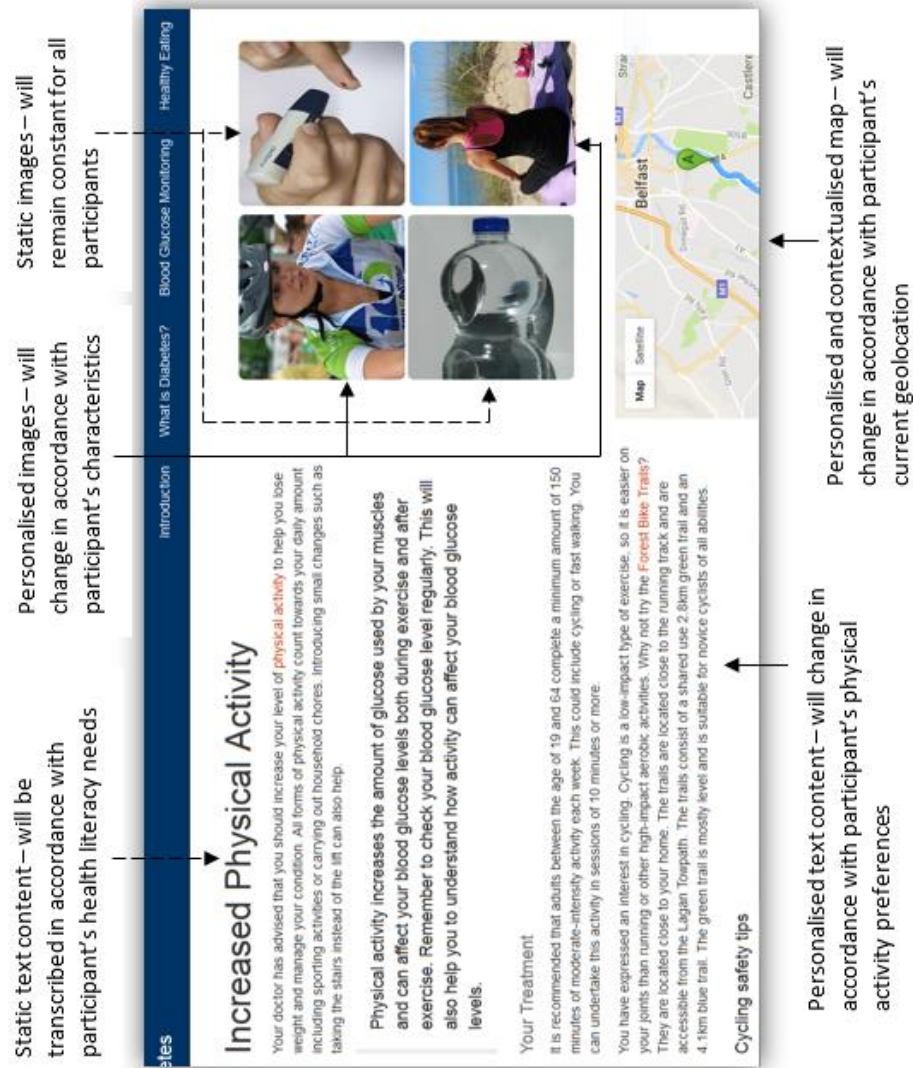


Fig. 6.6. Screen shot of education web page that focuses on physical activity, and contains static, personalised and contextualised content.

6.3.3 Eye Tracking

During an experiment a *Tobii X60 Eye Tracker* was used to unobtrusively record the eye movements of a participant. Fig. 6.7 illustrates the arrangement of the equipment during an experiment. A chair was placed in front of the eye tracker which a participant could adjust to a comfortable height and distance from the eye tracking device. The eye tracker uses near infrared illumination to detect the user's corneal reflection and in turn identify gaze points on the screen, at time intervals specified by the sampling data rate (60Hz) [458, 465]. Before the education was displayed each participant completed a calibration process by following a moving dot on the screen with their eyes. This enabled the eye tracker to accurately recognise the participant's gaze locations during the experiment. The collected data was transmitted to the *Tobii Studio* software application, where the *Tobii I-VT filter* [466] was applied to calculate various metrics.

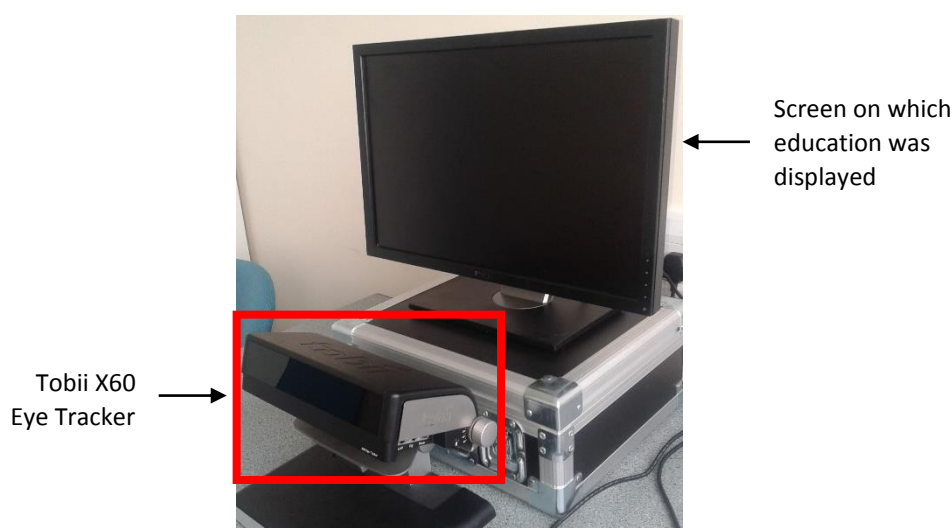


Fig. 6.7. Arrangement of Tobii X60 Eye Tracker and screen during an experiment.

6.3.4 User Engagement Scale (UES)

In the last stage of the experiment a participant completed the *User Engagement Scale* to report their engagement with the education. Related research had identified an association between perceived attractiveness of health information and an increased likelihood of attending to the information [467], thereby it was decided to focus on two dimensions of the participants' experience of the education; aesthetic appeal and

usability. Two of the UES subscales, *Aesthetics (AE)* and *Perceived Usability (PUs)* were adapted for use in the study. A copy of the two scales is available in Appendix E.

The AE scale focused on a participant's evaluation of the visual appearance of the education [455]. The scale comprised five items that addressed the following aspects of the education.

- The aesthetic appeal and attractiveness of the education.
- Whether the education was visually pleasing and appealed to the participant's visual senses.
- Whether the participant liked the graphics and images included.

Each item was scored on a five-point Likert scale of one to five (ranging from strongly disagree to strongly agree) with a final sum score within the range of five to 25. A higher end score indicated that the education was perceived as highly aesthetically appealing.

The PUs scale addressed a participant's emotional and cognitive responses to the education [455]. The original scale contained seven items, however it was decided that only six of the items were applicable to the patient education. Thereby the scale used in this study consisted of six items that were scored on a five-point Likert scale of one to five (ranging from strongly agree to strongly disagree) with a final sum score within the range of six to 30. A higher end score indicated that the education was perceived as highly usable. The six items focused on the following qualities of the education.

- Whether the education was confusing to use.
- Whether the participant felt frustrated, annoyed or discouraged while searching for information in the education.
- Whether the participant judged that searching for information in the education was mentally taxing or demanding.

6.3.5 Health Questions

After a participant had completed reading the education they were presented with four health questions, as listed in Table 6.1. The information required to answer the questions

was available in both education artefacts. Presenting a participant with health questions obliged them to search through the education. This ensured that the participant would have experience of seeking information within the education and could effectively evaluate the perceived usability of the education as measured by the items of the PUs scale (as described in Section 6.3.4).

Table 6.1. Health questions presented to a participant during an eye tracking experiment.

	Health Questions
1	Which test is commonly used to indicate a patient's blood glucose levels over the past 2-3 months?
2	"One common symptom of diabetes is unexplained weight loss." Is this statement true or false?
3	What is the recommended minimum amount of moderate intensity exercise that adults should complete each week?
4	Hypoglycaemia can occur if a patient's blood glucose level falls below 4mmol/l. Can you suggest 1 reason why hypoglycaemia may occur?

6.3.6 Statistical Analysis

Statistical analysis of the data was conducted using SPSS version 23. The mean and standard deviation of the AE and PUs scales were calculated through use of the descriptive statistics. The eye tracking metrics for each group were compared using a *t*-test, or a Mann-Whitney U test for data that was not normally distributed. Correlations between various eye tracking measures and the UES engagement scores were investigated through use of Pearson's product-moment correlation, Spearman's correlation coefficient or Kendall's Tau.

6.4 Results

All of the participants read the required portion of the education, and completed the four health questions and the UES. The mean experiment completion time for the experimental group was 738.56 seconds (SD 237.68), and the mean for the control group was 815.11 seconds (SD 250.47). A comparison of the mean completion times did not find a significant difference between the two groups ($t_{(16)} = -.665, p = .515$).

There was variation in the question scores achieved by the members of the experimental group; one participant (11.1%) achieved a score of two, while three participants (33.3%) achieved a score of three, and the remaining five participants (55.6%) answered all four health questions correctly. All of the members of the control group achieved a full score of four. Evaluation of the relationships between the health question scores and gaze behaviour found only one statistically significant relationship. For the experimental group a positive correlation was identified between question score and the total number of fixations on the text content ($r = .73, p < .05$). No relationships were identified between the question scores and the diabetes knowledge scores.

6.4.1 Engagement with the Generic Booklet and the Personalised Education

6.4.1.1 User Engagement Scale

The distribution of the aesthetics (AE scale) and perceived usability (PUs scale) scores for both formats of education are illustrated in Fig. 6.8. Analysis of the internal consistency of the AE scales found high reliability; for the personalised education scores $\alpha = .885$, and for the generic booklet scores $\alpha = .835$. For each participant, the five individual item scores were summed to provide an overall aesthetics score (out of 25) for the education. The scores for the personalised education ranged from 17 to 25, with a mean score of 21.2 (SD 3.31), and for the generic booklet the scores extended from 16 to 25, with a mean score of 20.56 (SD 3.08). Furthermore, t -test analysis did not identify a significant difference between the two sets of scores ($t_{(16)} = .442, p = .664$). The AE scale comprised five individual items which related to various attributes of engagement (as described in Section 6.3.4). The average scores for each of these items were calculated in order to ascertain if there were any differences in the participants' responses to the two education artefacts. For three of the items, aesthetic appeal, attractiveness and visually pleasing, the average score for the web-based personalised education was higher than the average for the generic booklet. However, for the items related to the graphics and images, and the appeal of the education the average scores were equivalent.

The internal consistency of the PUs scales was reliable; for the personalised education scores $\alpha = .871$, and for the generic booklet scores $\alpha = .867$. For each participant, the six individual item scores were summed to provide an overall perceived usability score

(out of 30). For the personalised education the perceived usability scores ranged from 22 to 30, with a mean score of 27.56 (SD 3.0), and for the generic booklet the mean score was 25.0 (SD 3.32) and the range was from 20 to 30. The two sets of scores were not significantly different ($U = 23.50$, $p = .136$). The average scores for the six PUs scale items (as described in Section 6.3.4) were calculated. The results indicated that for all six items (confused, annoyed, frustrated, discouraged, mentally taxing, demanding) the generic booklet had a reduced score when compared to the corresponding score for the personalised education.

Kendall's Tau was used to examine the relationship between the aesthetics and perceived usability scores. The results for both the generic booklet ($\tau = .030$, $p = .914$) and the personalised education ($\tau = .459$, $p = .114$) indicated that no statistically significant relationships existed between the scores for the two factors of engagement.

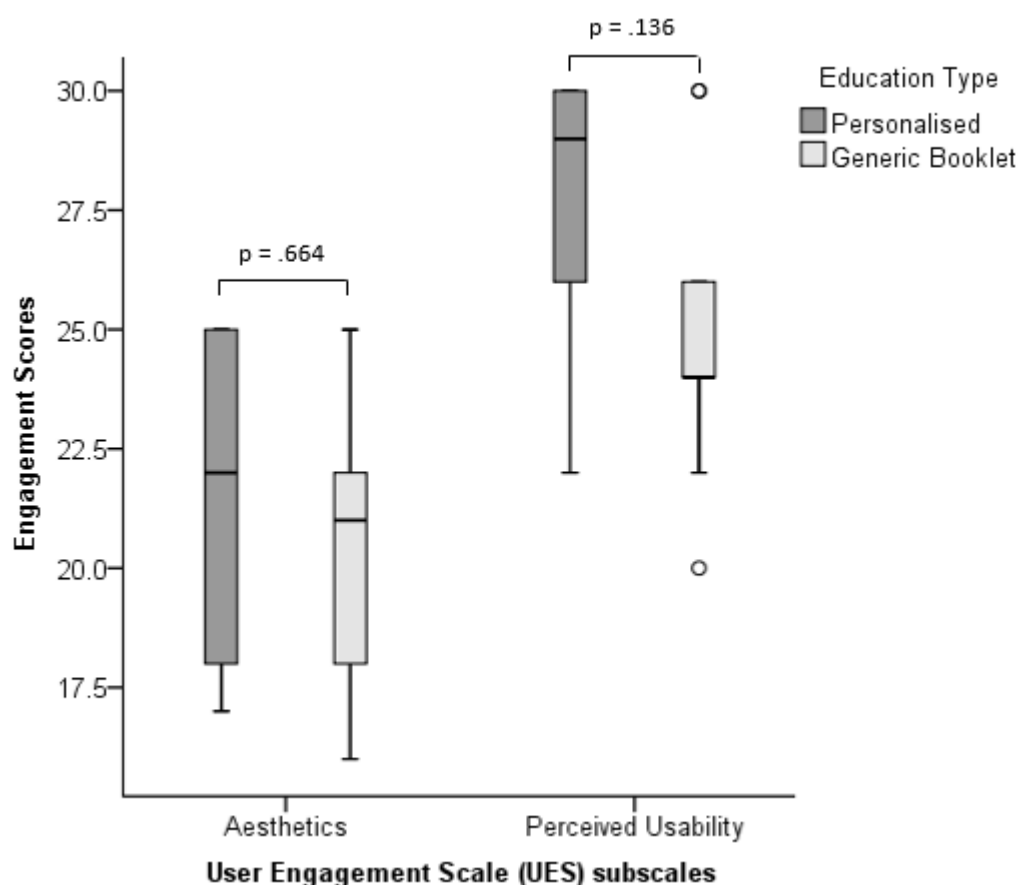


Fig. 6.8. Boxplots of the scores for aesthetics (AE scale) and perceived usability (PUs scale) for the generic booklet and web-based personalised education.

6.4.1.2 Gaze Behaviour

The generic booklet and personalised education comprised two types of educational content, text and images. However, while the images in the generic booklet would not alter throughout the experiment, for the personalised education the images were further categorised as, (1) *static images* which remained constant for all participants, and (2) *personalised images* that would change in accordance with the characteristics of a participant. The first stage of analysis focused on comparing the overall gaze behaviour of the members of the two groups during the experiment. *Tobii Studio* was used to define the text and image elements within the education as *Areas of Interest (AOI)*. Various fixation based metrics were then calculated for the AOI. The computed data was then exported so that further analysis could be completed in SPSS.

Fig. 6.9 illustrates the proportion of fixations for the different types of content in each education artefact. For both types of education the majority of fixations were directed at the text content. The percentage of fixations on image components was larger for the personalised education (6%) when compared to the booklet (2%), however this is not surprising as the personalised education contained a larger quantity of images. The booklet contained a larger total of static images ($n = 9$) than the personalised education ($n = 6$), however the proportion of fixations on the static images was equivalent (2%). The results of a Chi-Square test of independence indicated no statistically significant association between the format of education used and the number of fixations on the text or static images. While the personalised education contained a comparable amount of personalised ($n = 7$) and static images ($n = 6$), the percentage of fixations on the personalised images (4%) was double the proportion for the static images (2%).

The assessment of the overall gaze behaviour suggested that there may be a similarity between the two study groups in the overall proportion of attention directed towards static images. However, in order to clarify whether any similarities or differences existed in the engagement patterns it was decided to further evaluate the fixation behaviour of the two groups in relation to, (1) the static images, and (2) the text content of the educational artefacts.

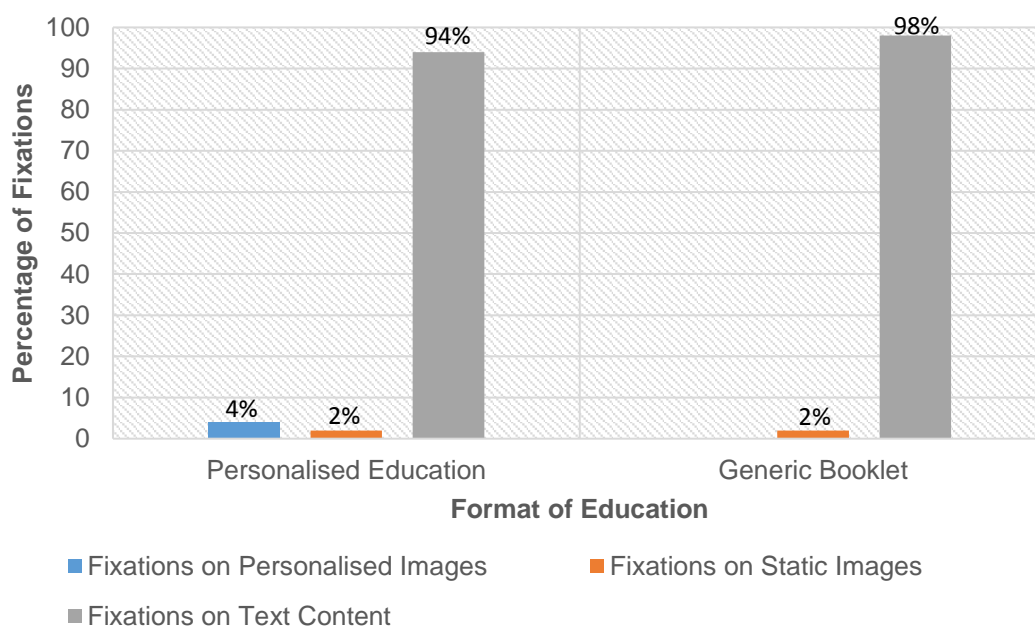


Fig. 6.9. Bar chart indicating percentage of total fixations on text, static images and personalised images for the personalised education, and percentage of total fixations on text and static images for the generic booklet.

6.4.1.3 Static Images

Fig. 6.10 indicates the distribution of the *total number of fixations* on the static images for the members of both groups. With an average of 32.56 fixations per participant (SD 16.76) the mean fixation rate for the control group was larger than the mean rate for the experimental group (19.0 fixations per participant, SD 12.67). T-test analysis did not find a statistically significant difference between the mean fixation count for the two groups, however the low p -value suggests a trend and perhaps greater sample size and statistical power would yield significance ($t_{(16)} = -1.94, p = .071$).

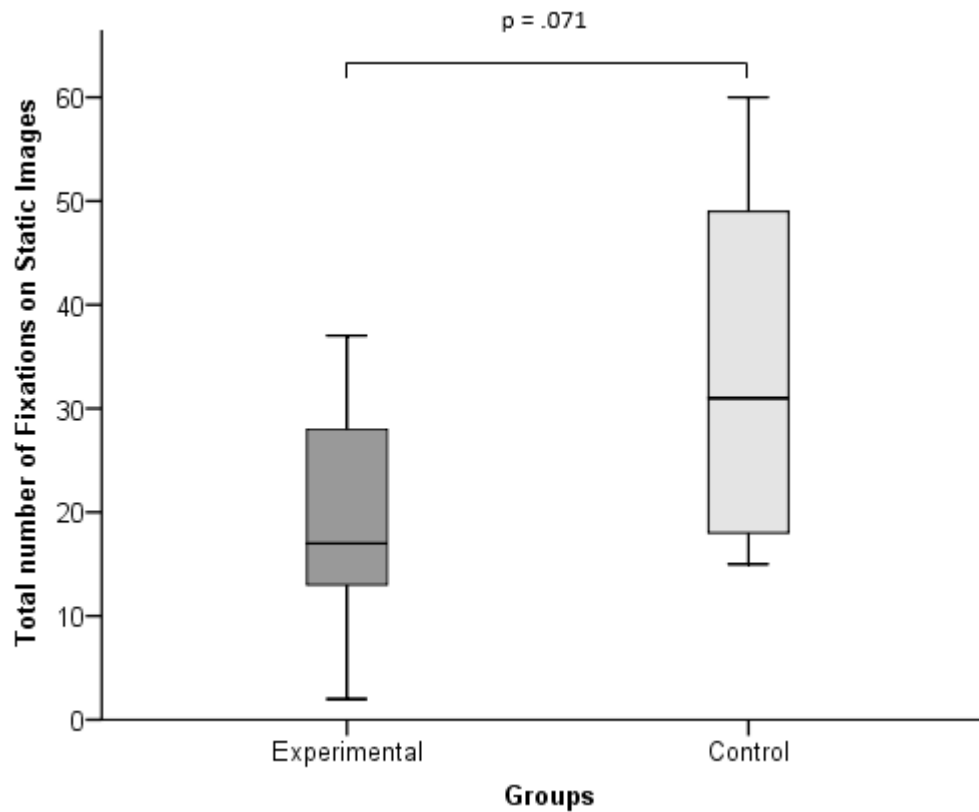


Fig. 6.10. Boxplots indicating total number of fixations on static images for experimental and control group.

It has been suggested that if elements of visual stimuli appear interesting, appealing, confusing or complex it may be more challenging for an individual to process these, and this may result in a longer period of fixation [458]. Evaluation of the fixation rate had suggested that the control group had attended more to the static images than the experimental group. However, the next stage of analysis attempted to quantify attention by using a temporal measure, *fixation duration*, which was defined as the length of a fixation in seconds. For each participant the duration of all fixations on each static image was summed. The distribution of *total fixation duration* on the static images for the members of both groups is indicated in Fig. 6.11. The average fixation duration per participant was larger for the control group (6.61 seconds, SD 4.18), when compared to the participants in the experimental group (4.1 seconds per participant, SD 3.53). However the mean fixation duration did not differ significantly between the two groups ($t_{(16)} = -1.38, p = .186$).

The metrics analysed to this point assess fixation on the static images. However, it is also useful to comprehend a participant's overall gaze behaviour on an AOI as this can reflect whether the participant returned to an AOI more than once while viewing the education. An AOI *visit* refers to a participant's observation of an AOI and includes both fixations and saccades. In Tobii Studio a *visit* to an AOI corresponds to the duration between the beginning of the first fixation on an AOI and the beginning of the first fixation outside that AOI. Evaluation of the visit data suggested that the participants in both groups did revisit the static images whilst reading the education. The mean number of visits to static images for a member of the experimental group was 12.67 visits (SD = 9.54), while for the control group the mean was much higher at 19.44 visits per participant (SD 10.2). However there was no significant difference between the average visits for the two groups ($t_{(16)} = -1.46, p = .165$).

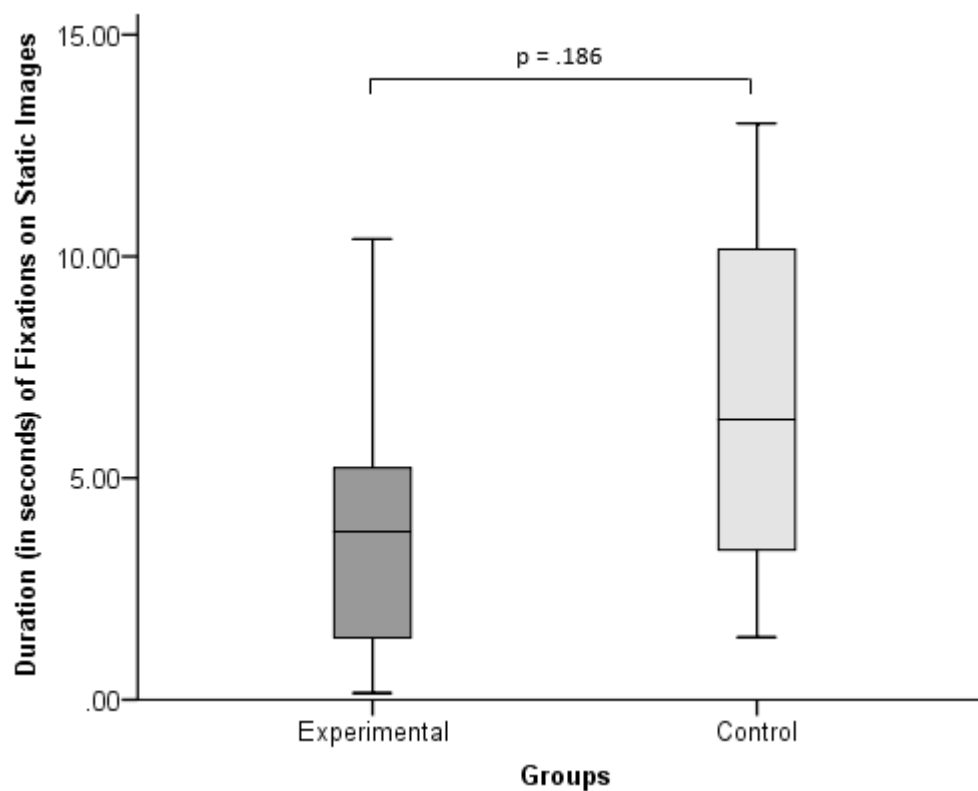


Fig. 6.11. Boxplots indicating total fixation times (in seconds) on static images for experimental and control group.

Relationships between engagement quantifiers

The eye tracking data signified the participants' level of interest in the static images, however it was decided to investigate how this data correlated with the participants' own appraisal of their engagement with the education (as denoted by the AE and PUs scores). Regarding the experimental group, no significant associations were found between the aesthetics scores and the perceived usability scores and any of the eye tracking metrics discussed. In contrast, for the control group statistically significant moderately positive relationships were identified between the aesthetics scores and the total number of fixations ($r = .808, p < .01$), total number of visits ($r = .793, p < .05$), and the total fixation duration ($r = .779, p < .05$) for the static images. However, no significant relationships were found between the perceived usability scores and the eye tracking data for this group.

6.4.1.4 Text Content

Fig. 6.9 indicates that for both study groups the majority of fixations attended to the text content of the education. Analysis of the eye tracking recordings suggested that most of the participants read through the text in a systematic manner and attended less to the images. For example, a gaze pattern for a member of the experimental group is illustrated in Fig. 6.12. The gaze plot (Fig. 6.12 (a)) illustrates the gaze path whilst viewing a section of a web page (what is diabetes) in the personalised education. The yellow dots indicate a fixation and are numbered in temporal sequence. A larger sized dot indicates a longer fixation length. Saccades are represented by yellow lines. A corresponding heat map (Fig. 6.12 (b)) is coloured coded to indicate the area which received most fixations (red), with reducing fixation amounts signified from yellow through green. Both figures illustrate how a larger amount of attention was directed towards the text in comparison to the images.

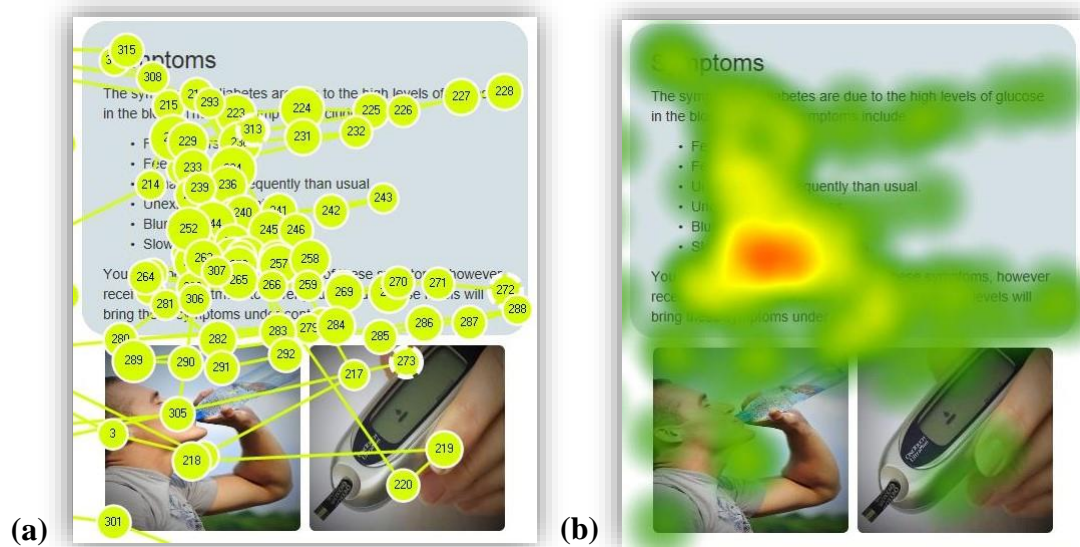


Fig. 6.12. Visualisation of fixations on personalisation education for a member of the experimental group, (a) gaze plot indicating gaze path, and (b) heat map indicating quantity of fixations on different education components.

The generic booklet had a greater volume of text, and correspondingly the mean time spent fixating on the text for members of the control group (352.1 seconds per participant, SD 183.86) was larger than the mean time interval for members of the experimental group (221.35 seconds per participant, SD 125.91). However a comparison of the data for the two groups did not find a statistically significant difference ($t_{(16)} = -1.76, p = .097$). The participants had indicated their diabetes expertise before the experiment began, thereby it was decided to investigate whether participants that had less knowledge attended more to the text information. Spearman's correlation coefficient was derived to determine whether any relationships existed, however no statistically significant correlations were found between diabetes expertise and total length of time spent fixating on the text information either for members of the control group ($r_s = -.101, p = .795$) or the experimental group ($r_s = -.463, p = .209$).

6.4.2 Gaze Pattern Analysis for the Experimental Group

6.4.2.1 Personalised and Static Images

A key objective of the study was to examine whether the participants that viewed the personalised education were attentive to the personalised features of the education. The

inclusion of images that were tailored to correspond with a participant's demographic characteristics comprised the most prominent visual personalised features. Thereby primary analysis focused on assessing whether there was a difference in the gaze behaviour for the personalised and static images. Fig. 6.13 depicts the distribution of the *total fixation rate* for the personalised and static images for the members of the experimental group. There was variation in the total number of fixations on the two types of images. The mean fixation rate for the personalised images (42.78 fixations per participant, SD 29.81) was more than double the mean rate for the static images (19.0 fixations per participant, SD 12.67). Furthermore, *t*-test analysis indicated that the fixation rate for the personalised images was significantly larger when compared with the rate for the static images ($t_{(8)} = 2.41, p < .05$).

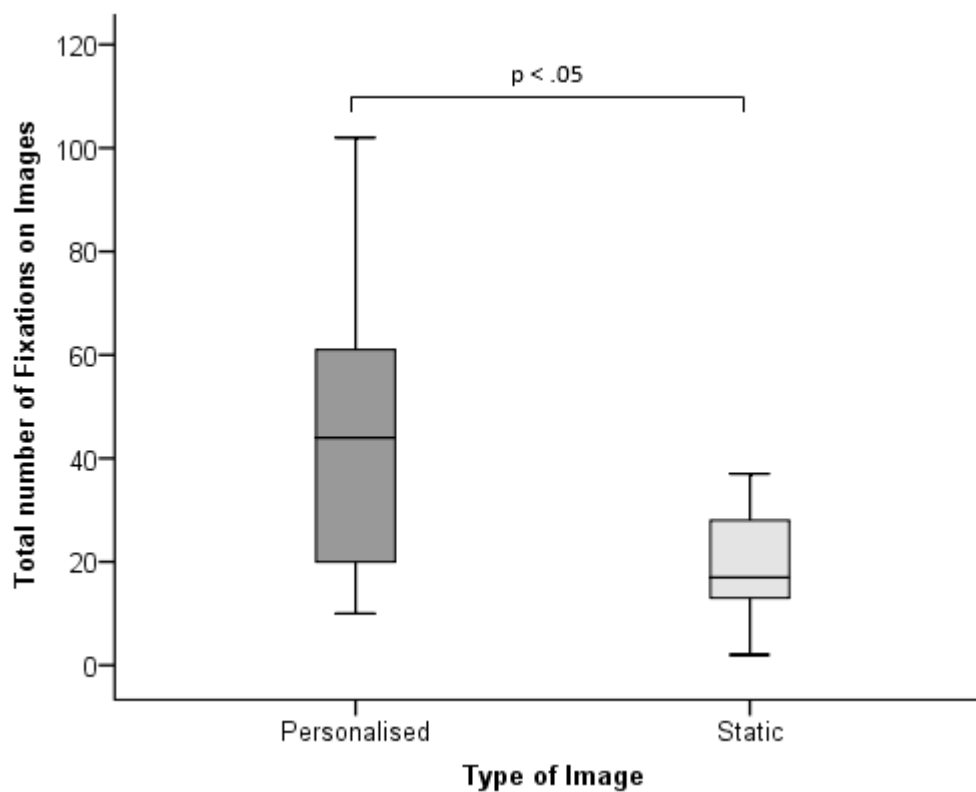


Fig. 6.13. Boxplots indicating total fixations on personalised and static images for experimental group.

The distribution of the *total fixation duration* for the personalised and static images is illustrated in Fig. 6.14. The average fixation duration on the personalised images was 8.58 seconds per participant (SD 6.67), which was twice the length of the mean fixation

duration on the static images (4.09 seconds per participant, SD 3.53). T-test analysis indicated that the difference between the duration means was very close to significance ($t_{(8)} = 2.29$, $p = .051$), however the result was not significant at the .05 level of significance. The mean number of *visits* per participant to the personalised images was 20.78 visits (SD 12.47), which was larger than the mean visit count for the static images (12.67 visits per participant, SD = 9.54). Moreover, this difference was statistically significant ($t_{(8)} = 2.36$, $p < .05$).

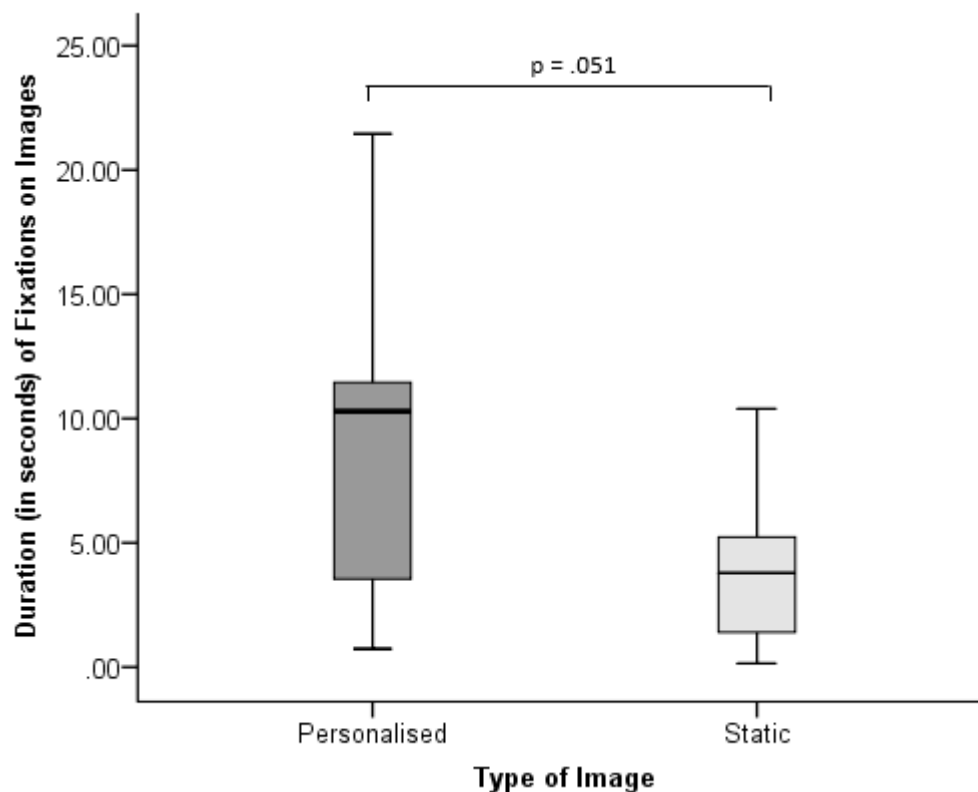


Fig. 6.14. Boxplots indicating total fixation times (in seconds) on personalised and static images for experimental group.

Correlations with the UES

As indicated in Section 6.4.1.3, for the experimental group, no statistically significant correlations were found between the aesthetics and perceived usability scores and the eye tracking metrics for the static images. Data analysis also examined the relationships between these engagement factors and the participants' level of attention for the personalised images. However, no statistically significant correlations were found between the scores for the aesthetics and perceived usability of the personalised

education and any of the eye tracking measures (fixation count, fixation length, visit count).

6.4.2.2 Personalised Components of the Web-Based Education

Evaluation of the fixation and visit data suggests that the participants that viewed the web-based personalised education were more attentive to the personalised images when compared to the static images. However, in order to further appraise whether some aspects of the personalisation were more effective in capturing the attention of the participants, fixation based metrics for each of the personalised images were examined. As described in Section 6.3.2.2 the personalised images were individualised by age and gender, however in the *physical activity* web page the personalised images were also contextualised to the participant's preferred sporting activity (as shown in Fig. 6.6). Table 6.2 indicates fixation based measures for the four images that were categorised as personalised and the two images that were classified as being both personalised and contextualised.

As indicated in the table, all except one of the personalised images received a higher quantity of fixations and were fixated on for a longer time period when compared to the personalised and contextualised images. The image with most fixations and the longest fixation duration (image no. 1) was positioned on the first web page of the education (introduction), and was the first image viewed by all the participants. This was also the largest sized image contained in the education (almost three times larger than the second largest image) and received 14% of the total fixations on the introduction web page. The second and third most attended images were also personalised images and appeared on the second (what is diabetes) and third (blood glucose monitoring) web pages of the education. The content of these three images contained an individual (male or female) whose face was clearly visible. The two personalised and contextualised images were positioned on the fifth and final web page of the education (physical activity), and the combined fixations for these two images comprised only 1.5% of the total fixations on this web page. Both images were characterised by containing an individual (male or female) involved in the participant's preferred sporting activity (e.g. jogging, running or cycling). The final web page (physical activity) also contained a Google map which illustrated directions from the participant's geolocation during the experiment to a sporting location. The map received the second highest quantity of fixations ($n = 119$

fixations) and in total the participants spent 23.72 seconds fixating on this image, which was the second longest fixation period for an image in the personalised education.

Table 6.2. Fixation metrics for the personalised images included in the web-based personalised education.

Web Page (with order number)	Image No	Image description	Personalised	Contextualised	Total number of fixations	Total fixation length (in seconds)	Average fixation length (in seconds)
(1) Introduction	1	Male/female, smiling	✓		120	25.27	0.21
(2) What is Diabetes?	2	Male/female, drinking water	✓		29	6.4	0.22
(3) Blood Glucose Monitoring	3	Male/female, smiling	✓		69	14.54	0.21
(4) Healthy Eating	4	Male/female with food, face not visible	✓		14	2.21	0.16
(5) Physical Activity	5	Male/female, sporting activity	✓	✓	17	2.68	0.16
	6	Male/female, sporting activity	✓	✓	17	2.4	0.14

The first web page of the education (introduction) included a greeting which contained the participant's name. This was included so as to encourage the participant to identify with the education. Analysis of the fixation data concluded that seven (77.8%) of the participants fixated on the greeting. The average length of a fixation on the greeting was 0.93 seconds, and this was larger than the average fixation length for the remaining text

(0.22 seconds per fixation) and the personalised image (0.21 seconds per fixation) presented on this page. However, the summed fixation lengths for the greeting comprised only 3.6% of the total fixation duration for this web page.

6.5 Discussion

Current approaches to patient education include the provision of standardised printed materials to all patients regardless of their demographics such as age and cognitive abilities. However, whilst these materials provide medically verified advice, the information provided may not be accessible for all patients. This suggests that an adaptable approach to providing health education may be beneficial for patients and physicians. Strategies to enhance the efficacy of health communications have included the tailoring of printed and electronic health interventions as appropriate for various demographic, behavioural and health attributes of health consumers [167, 170, 451]. Personalisation is a tailoring strategy which attempts to increase the self-relevance of information for an individual, and has been proposed as a means to focus attention on, and encourage central processing of health messages [165, 170, 467]. Chapter 4 detailed the technical development of a web-based framework that provides personalised education for diabetic patients. This chapter described a study which investigated whether users of the system would attend to the personalised features of the provided education.

6.5.1 Engagement Behaviour of the Control Group and Experimental Group Members

The first phase of analysis focused on evaluating engagement with generic and personalised diabetic education. Two groups of participants viewed either a generic booklet or web-based personalised education, and engagement patterns for the two groups were compared. The subjects completed the UES, a self-reporting tool through which they rated their experience for two factors of engagement, the aesthetics and perceived usability of the education. An unobtrusive eye tracking device was also used to capture the gaze behaviour of the participants as they read the education. Thereby both subjective ratings and objectively gathered data was assessed. The aesthetics scores for the two formats of education were comparable, however examination of the individual items that comprised the aesthetics scale identified that for three items, aesthetics appeal, attractiveness and visually pleasing, the personalised education

achieved a higher average rating. Despite differences between the two groups in the amount of time spent viewing the images, the average score for “liking” the images was equivalent. The average score for perceived usability was higher for the personalised education, however the scores did not differ significantly between the two types of education. The individual items that comprised the perceived usability scale related to the participants’ experiences of searching through the education and included emotions such as frustration, annoyance, discouragement and cognitive demand. For all six scale items the generic booklet received less positive ratings when compared to the scores for the personalised education. This suggests that searching for information in the generic booklet was more challenging in comparison with the personalised education. Nevertheless, despite the variance in the perceived usability scores all the participants in the control group correctly completed the four health questions, whilst only five members (55.6%) of the experimental group achieved a full score. It was unclear why there was a distinction in question scores between the two groups. The average diabetes knowledge score for the control group was greater than the mean score for the experimental group, and it is possible that previous knowledge may have contributed to the success of the experimental group. However, there was no significant difference between the mean scores for the two groups, thereby it is impossible to verify whether previous knowledge of diabetes had an impact on the health question scores for the control group.

Gaze behaviour metrics were used to identify where the visual attention of the participants was located on the education. For the members of both groups most attention was focused on the text content. A comparison of fixation on the text content suggested that the members of the control group spent a longer period of time reading the text when compared to the experimental group. However, the generic booklet contained a larger volume of text which may have contributed to the difference in fixation time. There are several reasons why both groups may have spent such a large proportion of time attending to the text. If visual information requires increased cognitive processing then an individual may spend an increased amount of time looking at the information [458, 462]. The ratings for diabetes knowledge indicated that 15 participants (83.3%) had rated their knowledge in the lower half of the scale, thereby it is likely that reading new information and encountering unfamiliar terminology may have resulted in increased attention. However, no relationships were found between

level of diabetes expertise and amount fixated on the text. Moreover, it is also probable that the subjects spent a longer period of time reading the text content as they were aware that they would be answering questions about the information in the second stage of the experiment.

Appraisal of overall gaze behaviour indicated that the percentage of fixations on static images was equivalent for both groups. However, further examination of the data indicated that the participants that viewed the generic booklet fixated on the static images more frequently and for longer time intervals when compared to the participants that viewed the personalised education. Moreover, the members of the control group tended to revisit the static images more frequently than members of the experimental group. This behaviour may have been due to the types of images included in the booklet. The images in the generic booklet included a full-page diagram which explained the function of insulin. The eye tracking data suggests that this diagram was more successful than any other image in the booklet in capturing the attention of the participants (receiving 55% of all fixations). Furthermore, it was also fixated on more frequently than any of the static images in the personalised education. The diagram can be classified as a *cognitive illustration*, an explanatory image which complements information and assists readers to more fully understand the text content [468, 469]. The remaining images in the booklet and all the static images in the personalised education are *affective illustrations*. These are illustrations that are used to draw attention to information, augment the attractiveness of the content and enhance the enjoyment of the information for the reader [468, 469]. Cognitive illustrations have been found to attract more attention than other illustrations in some online health information materials [468]. Similarly, in this study the diagram was especially successful in capturing the interest of the control group. Hence, it is probable that the increased level of interest for the diagram contributed to the increased attention towards the static images exhibited by the control group. Correlation evaluation highlighted an association between increased fixation on the static images by members of the control group and increased aesthetics scores for the generic booklet. However, a similar relationship was not identified for the subjects that viewed the personalised education.

6.5.2 Attention Towards the Web-Based Personalised Education

The second stage of analysis concentrated on the participants that viewed the web-based personalised education. Viewing behaviour suggested that these participants looked at the personalised images more often and for double the length of time than was spent fixating on the static images. The personalised images had been matched to the participant's gender and age group, and there are several reasons why the personalised images derived such increased notice when compared to the static images. A recent study of engagement with images on Instagram found that photos containing faces were more engaging, receiving more likes and comments, in comparison to photos not containing a face [470]. Thereby as most of the personalised image content represented people, while the static image content focused on medical devices and food items, this may have encouraged engagement with the personalised images. Moreover, it is often found that humans have a preference for focusing their attention on people within pictures and illustrations. For example, during an experiment which presented image pairs comprising a person-present and person-absent scene, the subjects fixated on the person-present scene first, and for a longer period of time [471]. Participants also have an increased likelihood to quickly locate and fixate on people when depicted in social scenes [472, 473]. Furthermore, several eye tracking studies have identified attentional bias towards human faces in pictures, with a particular focus on the eyes [472-476]. It has been proposed that enhanced interest in the eyes is due to a tendency for humans to try to infer social information from an individual's eye gaze, and thereby better comprehend the meaning of a scene [472, 477]. It is evident that when viewing images, humans tend to focus their attention on meaningful objects such as other persons. Thereby, as all of the personalised images contained people, this may have naturally attracted the attention of the participants. Consequently, it is probable that this contributed towards the increased level of focus towards the personalised images.

A final reason for the increased attention towards the personalised images could be attributable to the composition of the experimental group. Many of the participants were from a Computer Science background and thereby may have been naturally interested in the concept of computerised personalised education. Consequently, the increased attention towards the personalised images may have been due to an appreciation of the technical challenge involved with creating online personalised education.

Examination of the gaze metrics for the personalised images also highlighted some variations in attention to the personalised images. The fixation rate and length suggested that the images positioned on the first three web pages were the most frequently attended to. The image located on the first web page (introduction) was the first image attended to by all the participants on reading the education. This image had the largest quantity of fixations and longest fixation length. However, this image was larger than all other images in the education, thereby it is possible that the size and location of this image contributed to its increased focus. The images situated on the *physical activity* web page received less attention than images placed on earlier web pages. This was surprising as the content focus of these images matched the participant's age group, gender and preferred sporting activity. However, for all the participants the *physical activity* web page was the last web page read as they progressed through the personalised education, thereby participant fatigue may have contributed to the lower fixation rate. Nevertheless, the Google map which was also positioned on the *physical activity* web page received almost as many fixations as the image on the first web page. It is likely that the subjects found the map to be more cognitively engaging as it included a mixture of content (text and illustration) which conveyed information that focused on their immediate geolocation. This suggests that providing information relevant to the context in which a user currently resides may capture the user's attention.

The personalised greeting was not attended to by all the participants. This suggests that identifying a user by name within health information may have variable success in capturing the attention of the reader. However, it is also likely that some participants may have simply overlooked the greeting as they were eager to start reading the diabetes information, which they may have perceived as more relevant to the experiment.

The performance of both study groups suggests a high level of engagement amongst the participants. However, it must also be acknowledged that the experimental conditions may have had an impact on the participants' behaviour. As detailed in Chapter 3, the behaviour of a participant may alter during an experiment due to an awareness of being observed rather than the actions of the investigator, a situation known as the *Hawthorne effect* [382, 383]. Thereby, it is possible that the attendance of the investigator may have influenced the motivation of some subjects during this study. For example, due to the presence of the investigator, a participant may have felt obliged to complete all stages

of the experiment. One consequence of recruiting within a university was that a number of the subjects were from a Computer Science background and were colleagues of the investigator. Subsequently these participants may have been eager to perform well in the experiment, thereby encouraging their commitment to the tasks. Moreover, as previously described, participants from a Computer Science background may have been motivated by an interest in the concept of personalisation, which increased their attention to the personalised content. One aspect of the study which may have positively impacted the commitment of the participants was the use of the eye tracking equipment. Some of the participants were intrigued by the novelty of eye tracking and were keen to participate in the study. However, it was hoped that providing interesting medical based information would distract the participants' attention from the eye tracking equipment throughout the experiment.

6.5.3 Study Limitations

There were a number of limitations with this study. The primary limitation was related to the reduced size of the study groups. The volume of eye tracking data collected and the UES based data did provide an adequate amount of data to complete statistical analysis of various attributes of engagement. Moreover, statistically significant and non-significant relationships between engagement attributes were identified for the study population. However, an increased sample size may have highlighted further patterns in the gaze behaviour of the participants, or may have increased the statistical significance of the identified relationships. The second limitation relates to the reading behaviour of the participants during the study. All the participants read the full portion of the education as requested, however as the study was performed under laboratory conditions it is likely that they may have completed reading the education in less time than would be taken if viewing information for a genuine health issue. However, examination of the eye tracking recordings indicated that the participants did progress systematically through the education artefacts, which enabled attention to the different components of the education to be analysed.

6.6 Conclusion

The research question queried whether any relationships exist between the format of education being used and the engagement behaviour exhibited by an individual. To

answer this question a study was designed which investigated characteristics of engagement for participants that used either a generic education booklet or web-based personalised education. A comparison of overall viewing behaviour determined some similarities between the two groups as it was established that most of the participants' attention was directed towards the text information and less attention was directed towards the images. Moreover, with regard to the participants' experience of the usability and visual appeal of the education no significant difference was found between the two formats of education. Statistical evaluation indicated that attention towards the static images was positively associated with the aesthetic appeal of the generic booklet. This would suggest that for some individuals the inclusion of images may contribute to the visual appeal of health education. However, for both forms of education no associations were found between the usability of the education and the amount of attention directed towards either text and images components.

Analysis of the gaze behaviour of the participants that viewed the personalised education indicated that the attention directed towards the personalised images was twice the amount than was directed towards the static images. Thereby, personalisation of the image content was successful in attracting the attention of the participants. Variations within the gaze behaviour revealed that the participants attended more to images which focused on an individual's face. Nevertheless, it must also be considered that the volume of fixations on the personalised images tended to reduce as the participant progressed through the education. Thereby it is possible that participant fatigue may have impacted on the amount of attention paid towards images that appeared later in the education. There was one distinction among this pattern which related to the personalised map. Despite being the last image viewed as the participants read through the education, the map attracted the second highest quantity of fixations. This suggests that the inclusion of visual information that is personalised and contextualised to the current geolocation of a health consumer may be effective in attracting attention. The inclusion of the participant's name had variable success in attracting the attention of the participants.

With relation to the research question, this results of this study suggest that there are similarities and distinctions in the manner in which health consumers use different formats of health education. The viewing behaviour indicated that whichever format of

education was used, the text content was the primary focus for the participants. This underlines that text content is the central element of a health education artefact for many users and draws most attention. Thereby one approach to improving the usefulness of health education is to ensure that the readability of the text information is comprehensible for the target audience, and focuses on their health information needs. For some of the participants an association was established between aesthetic engagement with the material and the inclusion of images. This indicates that for some individuals the inclusion of images in health education may help to motivate engagement with the material. Despite a lack of statistically significant relationships between the eye tracking metrics and the factors of engagement, the participants' gaze behaviour indicates that the personalised features did attract the attention of the participants. The viewing behaviour suggests that the inclusion of image content that focuses on persons that a user can identify with may attract the attention of the user. Moreover, the participants' visual reactions to the contextualised image map also suggests that providing information that is relevant to the current context of an individual may also motivate interest. The results indicate that image content can successfully attract the interest of a reader, and this provides an opportunity to improve the efficacy of patient education. Including images that reflect the characteristics and context of the user may encourage increased attention and assist users to identify and engage with the material. Moreover, utilising tailored images to complement the educational focus of the material may assist to inform the user, and enhance the efficacy of the education.

Chapter 7

Conclusion and Scope for Future Work

7.1 Introduction

Chapter 1 described the context of this thesis by highlighting the necessity for citizens to have access to accurate and coherent health information so that they can better understand and fulfil their healthcare needs. However, the channels from which individuals are consuming information are diversifying. This is mainly due to the increased availability of the Internet, which has enabled access to a varied choice of online health information resources. Nevertheless, engaging with health information can be challenging for many people, due to their own personal skills or the characteristics of the health materials obtained. This PhD addressed several issues associated with the usage of health information. This included a focus on the human skills and behaviours associated with obtaining and engaging with health information. Furthermore, the research also investigated the use of semantic web technologies to facilitate the provision of personalised electronic health information.

The thesis began by indicating the aim and objectives of this PhD project, and proposing several contributions to knowledge that would result from the research. Earlier chapters have described the research projects and technical solutions that have been completed in fulfilment of this PhD. This final chapter will explicate how the objectives were achieved through completion of these studies and the resultant knowledge contributions. Moreover, the discussion will also address several limitations inherent in the research studies, and will detail the scope for future work.

7.2 Research Objectives

The research objectives for this PhD have been fulfilled through the completion of the four research studies described in chapters 3, 4, 5 and 6, and a literature review as detailed in Chapter 2. The activities carried out to achieve each objective are detailed below.

Objective 1.

To develop an understanding of current research activities focused on the communication channels, strategies and technologies used in the provision of electronic health information.

This objective was achieved through a comprehensive appraisal of associated literature, as reported in Chapter 2. The review began by appraising the methods by which individuals obtain and use health information from various sources, and the associated challenges. The review then concentrated on the use of personalisation in healthcare applications, and highlighted several technical difficulties that are associated with delivering personalised and contextualised services, including user modelling. Subsequently, the increasing use of ontologies in biomedical and healthcare domains was discussed, along with recent research efforts related to ontology engineering. The literature review identified several areas for further research in the sphere of online health information seeking, personalised health information provision and ontological modelling.

Objective 2.

To use software tracking and log analysis to classify and quantify the behaviours of health information seekers when searching for and using online health information, and assess whether these behaviours are associated with health literacy or eHealth literacy skills.

This objective was fulfilled by the completion of a study which aimed to investigate the online health information seeking behaviours of health consumers. This study is described in Chapter 3. The study, which involved 60 participants, set online searching tasks and then monitored the activities of the participants as they used online searching tools and health information resources to answer health questions. The health literacy and eHealth literacy skills of the subjects were assessed using validated instruments. A data set with 360 records was collected from the study, and a bespoke electronic tool was used to parse the data and classify and quantify various behaviours. The behaviours evaluated included the time spent on different types of resources and classification of the resources used. A comprehensive set of statistical tests were executed to critique the associations between the online health information seeking behaviours and health literacy and eHealth literacy. In addition, various commonly used search strategies were

identified and evaluated, such as the activities associated with using search engines to find health information. The analysis also considered various contextual factors which may have affected the performance of the participants.

Objective 3.

To design and implement a web-based technical solution that will provide electronic patient education that is tailored as suitable for the personal health needs and skills of patients.

This objective was achieved through the development of a web-based architecture that created tailored education for diabetic patients. The design, implementation and testing of this architecture was detailed in Chapter 4. A key element of the architecture comprised an ontological knowledge base which included a model of several clinical aspects of diabetes, and a user model representation of a diabetic patient. A rationale for choosing an ontology as the knowledge base and selecting OWL as the representation language was provided, alongside a detailed account of the various stages of ontology engineering. The implementation of a SWRL personalisation rule-set was also detailed. These two components function to resolve the personalised features that will be included in the education, as applicable to the personal health needs and skills of each patient. Testing verified that the system produced accurately tailored education. For each patient, the educational content was tailored to focus directly on the symptoms, treatments and complications as relevant to the patient's health experience, and the text content was transcribed at a readability level appropriate for a patient's health literacy skills. The image content was personalised to reflect the patient's age and gender, and the layout was adapted as suitable for their level of digital experience.

Objective 4.

To develop a novel approach for collaborative evaluation of an ontology in the health domain, and evaluate whether this approach is accessible for ontology engineering novices.

This objective was fulfilled through the completion of a study which aimed to evaluate whether individuals with limited ontology modelling experience could complete ontology engineering tasks. The study involved 21 participants, and is described in Chapter 5. The methodology comprised a set of ontology engineering tasks which

collected domain knowledge. The presentation of the tasks was formulated to assist the participants with transcribing domain knowledge. Furthermore, informational artefacts, including a multimedia presentation, were provided to describe ontology engineering activities and assist with modelling decisions. The knowledge contributions were used to evaluate the patient education ontology. In addition, an analysis was also completed to appraise which types of tasks were most accessible and most challenging for the participants.

Objective 5.

To appraise the comprehension and engagement behaviours exhibited by individuals when interpreting generic patient education material and computer-aided personalised patient education material, and evaluate whether there are any similarities or distinctions in engagement patterns between both approaches.

This objective was fulfilled through the completion of a between subjects study which aimed to evaluate the engagement behaviours of participants when using either a generic patient education booklet or electronically derived personalised patient education. The study, which involved 18 participants, is described in Chapter 6. The methodology was designed to gather both objective and subjective measures of engagement. During an experiment the eye movements of a participant were unobtrusively recorded as they read through and searched for information in one of the artefacts. This enabled the collection of natural visual responses to the education. Corresponding attributes of gaze behaviour, such as the amount of time spent looking at visual stimuli, were quantified using eye tracking software. In order to appraise individual perceptions of the education, each participant completed the UES. This provided subjective measures of engagement for each participant. Statistical analysis was used to evaluate the engagement behaviours of each study group in reaction to similar features within the two materials, and identify comparable and distinctive patterns. This afforded a thorough, and statistically substantiated comparative analysis of the engagement patterns exhibited by the participants when using the different formats of patient education.

Objective 6.

To assess whether the engagement behaviours exhibited by individuals will vary in response to viewing different aspects of personalisation included in a patient education artefact.

A second goal of the study described in Chapter 6 was to determine whether there was variability in the engagement behaviours displayed when viewing different personalised features in patient education. This phase of analysis focused on the engagement patterns of the participants that viewed the electronic personalised patient education. Traits of engagement such as the time spent viewing different personalised elements and the total quantity of fixations upon these features were measured. This enabled the identification of fluctuations in engagement levels between different personalised features. Contextual factors, such as the positioning of the features in the education, were also considered as explanatory factors.

7.3 Summary of Knowledge Contributions

The research completed during this PhD has resulted in various contributions to knowledge as detailed below.

1.1 The design and implementation of a methodology to track and quantify the online health information seeking behaviours, and measure the health literacy and eHealth literacy skills of health information seekers. (Objective 2)

Chapter 3 presented a methodology that was used to monitor and quantify the online health information searching behaviours of the participants, and to measure their health literacy and eHealth literacy skills. The methodology included online searching tasks, the use of electronic tools to record health information seeking activities, and screening instruments to measure the participants' skills. Six health questions were devised, and during an experiment these were presented to a participant through a Mozilla Firefox extension, the HCI Browser. This tool automatically recorded a participant's interactions with the web browser as they searched for information to answer a health question. Screen capture software was also used to record searching activities that were not logged by the HCI Browser. A bespoke parsing tool was used to parse the HCI Browser log files, and quantify various online health information seeking behaviours.

The parser produced a spreadsheet containing behavioural data for each participant. The methodology also included the use of two validated instruments to provide a quantitative measure of the participants' skills. The NVS was used to appraise a participant's level of health literacy, and the eHEALS was used to quantify eHealth literacy skills.

This methodology encompasses a contribution to knowledge as it provides a method for collecting, quantifying and analysing online health information seeking behaviours and measuring health literacy and eHealth literacy. The methodology was validated in a study of online health information seeking, and supported a comprehensive evaluation of the relationships between a range of online health information seeking behaviours and health literacy and eHealth literacy skills. The use of health question search tasks, the HCI Browser extension and screen capture software enabled a sizeable data set to be collected which incorporated a variety of online health information seeking behaviours. The parsing tool provided human readable records of searching behaviours in a format that can be used with statistical analysis software. This tool has been made available online for use by other researchers (refer to [77]). The methodology also included verified instruments which enabled accurate measurement of the health literacy and eHealth literacy skills of the participants. The methodology is easily replicable and extendable, and could be utilised for similar studies in this area.

1.2 The implementation of an extensible web-based architecture, comprising an ontological knowledge base and rule-based personalisation, which will deliver education that has been tailored to the preferences, personal health needs, health literacy and eHealth literacy competencies of individual patients. (Objective 3)

A major research output from this PhD is an extensible web-based architecture that provides patient education that is personalised to the preferences, health needs, and health literacy and eHealth literacy skills of diabetic patients. The architecture comprises an OWL ontology knowledge base and a SWRL rule personalisation component. The ontology incorporates a model of clinical knowledge related to diabetes, and a user model which represents the personal characteristics, health status, educational abilities and preferences of a diabetic patient. The SWRL rules establish the

conditions for personalisation, and a Pellet reasoner is used to reason with the rules and deduce the knowledge required to personalise the education. The architecture includes various Java-based logical modules and a database in which the educational content is persisted. The communication infrastructure is implemented using the Spring MVC framework, and the education is constructed as a JavaServer Page and is presented in a web browser.

The patient education architecture provides contributions to knowledge in the areas of personalisation and health education. Primarily the system comprises a fully functioning technical solution for the provision of personalised education for diabetic patients. The education incorporates several personalised features including tailoring of the content as applicable for each patient's particular health information needs and health literacy skills. The visual layout and typography of the education is presented as suitable to a patient's preferences, visual acuity and eHealth literacy skills. Various other personalised elements are included so as to assist the patient to identify and engage with the education. This includes identifying the patient by their name and the inclusion of images that have been matched to the patient's age and gender.

The system also makes contributions to semantic web technology research. The architecture comprises a layered structure which incorporates several technical components that are easily extendable or re-usable. For example, the OWL ontology provides a user model which represents a wide range of personal, health and educational characteristics of a patient, and also captures a model of the semantics of diabetes related clinical data. The ontology could be re-used in other semantic web based systems that require to represent knowledge in this area. In addition, the architecture was designed to facilitate further adaptation of the education. For example, the SWRL rule-set was devised to be easily extendable to increase the level of personalisation in the education by adding new personalised features.

1.3 The design and implementation of a method for collaborative ontology evaluation that supports domain experts with variable modelling experience to complete ontology engineering tasks and contribute domain knowledge. (Objective 4)

Chapter 5 described a methodology that was formulated to engage ontology engineering novices with a range of ontology engineering tasks. The domain knowledge collected through these tasks was used to evaluate the patient education ontology. The methodology included informational artefacts that described ontology engineering activities, and an intuitive user interface to assist with transcribing a modelling decision. In the first stage a multimedia presentation was provided to introduce various ontology modelling practises and associated terminology. The ontology engineering tasks were presented as an online survey. Five types of engineering activities were incorporated into the survey including proposing new class properties and structuring data in a subsumption relationship. Moreover, information relevant to a modelling choice such as health-related descriptions or ontological diagrams were also provided. These features assisted the participants with completing a range of the survey tasks.

This methodology comprises a contribution to knowledge as it provides a method for assisting domain experts with limited modelling experience to become involved in ontology evaluation processes. The methodology incorporated explanatory information in various media formats which provided sufficient knowledge to enable the participants to complete various engineering tasks. The presentation of the engineering tasks also assisted the participants with submitting modelling decisions. The methodology was employed in a study which aimed to evaluate whether individuals with limited technical modelling experience can effectively contribute to the evaluation of an ontology. The participants' contributions were utilised to evaluate various models within the ontology, and to expand the scope of the ontology by adding new knowledge. Thereby, the methodology provides a means to expand ontology evaluation practices to include individuals that have limited experience of modelling ontologies or other technical artefacts.

1.4 The design and implementation of a methodology to record subjective and objective measures of engagement with patient education artefacts. (Objectives 5 & 6)

Chapter 6 presented a methodology for a between subjects study that was used to record user engagement with generic and personalised patient education artefacts. The methodology was formulated to document both subjective and objective quantifiers of engagement. During an experiment a participant read either a diabetic education booklet from a UK charity, or personalised education as created by the patient education architecture. Following this they received four health questions and had to search for relevant information within the education. An eye tracker was used to record the visual responses of the participants as they used the education, and eye tracking software calculated various characteristics of eye movements. This enabled the collection of objective indicators of engagement. The participant then completed the UES, a validated self-report instrument through which they indicated their perceptions of the aesthetic appeal and perceived usability of the education. This enabled a participant's subjective responses to the education to be quantified. Statistical analysis was used to uncover relationships between the UES data and eye tracking movements, and to uncover similarities and differences between the engagement behaviours of the two groups of participants.

This methodology affords a contribution to knowledge as it presents a method for gathering and analysing distinct quantifiers of a user's engagement with patient education. The use of an unobtrusive eye tracker device to automatically capture the reader's gaze behaviour facilitated the recording of natural responses to the education. Furthermore, using the UES enabled the users to deliver their subjective interpretations of the visual attractiveness and usefulness of the education. The methodology was validated in a study which appraised the engagement behaviours of participants that viewed either a generic or a personalised patient education artefact. Statistical analysis identified analogous and dissimilar patterns of engagement between the two participant groups. It was also found that there can be variability in user engagement with different personalised features in a patient education artefact. This methodology could be easily replicated in similar research studies which require to measure multiple attributes of engagement with health information.

7.4 Limitations

The most prominent limitations of this PhD relate to the composition and size of the study populations for several of the research studies. These limitations are detailed below.

- Chapter 3 described a study which aimed to investigate online health information seeking behaviour. Sixty participants were engaged to take part in the study, however there was a lack of variety within the study population with relation to age and educational attainment. The recruitment phase focused on attracting subjects from within a university campus. Therefore, although there was variation in the background disciplines of the subjects, approximately 62% had achieved an undergraduate or postgraduate degree. With relation to the ages of the participants the range was from 18 to 59, however the average age of a participant was 27 years. Consequently, the study population could be characterised as being younger and highly educated. As discussed in Chapter 3, the subjects exhibited a range of online health information seeking behaviours which enabled a comprehensive critique of their performance. However, engaging a larger study population, with increased diversity in terms of age and qualifications, may have broadened and enriched the study findings. For example, older individuals are increasingly using the Internet to seek health information [28], thereby the inclusion of older people in the study may have uncovered online searching strategies that are comparable or dissimilar with a younger population group. Moreover, analysing the health literacy and eHealth literacy competencies of an older age group, and associated searching activities, may have strengthened the significance of the study findings.
- Some of the more unexpected results in Chapter 3 related to the weak correlations between the eHEALS scores, the NVS scores and the health question scores. Although eHEALS has been repeatedly validated as a reliable measure of eHealth literacy skills [351, 360, 478, 479], it did not appear to accurately reflect the online health information seeking skills of all the participants. Several possible reasons for this finding were proposed, including the overestimation of eHealth literacy proficiency by the participants. Furthermore, due to the context of the information seeking (an experiment) some of the subjects may not have exhibited their full skill

set to locate and use online health information. However, as the study aimed to specifically investigate whether eHealth literacy could be associated with an individual's OHISB, the limitations of the eHEALS tool were disappointing. A different approach to measuring eHealth literacy may have provided a more accurate estimation of the eHealth literacy skills of the study population. However, whilst several tools for measuring eHealth literacy have since been proposed [384, 385], unfortunately these were not available when the study was taking place.

- Chapter 6 described a study of engagement with patient education, however the size of the study population was reduced as only 18 individuals participated. Furthermore, there was a notable representation of highly educated persons in the study as 83% had achieved an undergraduate or postgraduate degree. Several related studies have identified an association between the manner in which individuals read and process health information and literacy level [57, 58, 461]. Thereby, having a larger and more educationally diverse study population may have augmented the study results, for example, by enabling an analysis of whether engagement behaviours could be associated with education level.
- Chapter 5 detailed a study which aimed to include domain experts in the evaluation of the patient education ontology. Whilst the size of the study population was limited (21 subjects), a main constraint of the study relates to the involvement of only two participants from a health-related discipline. Although a notable amount of data was collected from the tasks, the involvement of a greater number of health domain experts may have increased the variety of the data collected and enabled more diabetes specific terminology to be gathered. Adding domain knowledge of this type may have improved the scope and quality of the ontology.

7.5 Future Developments

Appraisal of the research studies and associated contributions to knowledge has suggested several areas for enhancement. These proposals are detailed below.

Online Health Information Seeking Behaviours (Chapter 3)

- During this study a substantial amount of analysis focused on the manner in which the participants read through search engine results lists and chose a link to follow. Examination of these behaviours required the investigator to scrutinise screen recordings and determine how far the participant had scrolled through the results list, which link the participant chose and its position in the list. One method by which to clarify which search results the participant focused on before making a decision would be to introduce eye tracking into an experiment. By tracking the eye movements of a participant it could be determined the order in which the participant read the results, the intensity of interest for each link, and whether they focused more intensely on the URL, the title or the snippet before moving on or selecting this result. Collecting this data may help to elucidate which aspects of information a health consumer prioritises when choosing to follow or disregard a particular result in the list.
- Throughout the study the participants used a range of SERPs and accredited and unaccredited health resources to gather information. However, although it was possible to quantify and evaluate this behaviour, it was impossible to clarify the particular reasons why the participants utilised these resources. Introducing the *think-aloud* method may encourage a participant to comment on their actions during an experiment and provide valuable feedback regarding their intentions. This would assist an investigator to comprehend a participant's reasoning behind using a particular health resource, as well as indicating their intentions when carrying out other searching activities.

User Engagement with Patient Education (Chapter 6)

- This study utilised an eye tracker to record a participant's eye movements as they read patient education, and employed the UES tool to gather a participant's subjective perceptions of the materials. These provided different measures of user engagement and enabled a comprehensive critique of engagement behaviours. However, user engagement is a multifaceted concept which is challenging to quantify. Accordingly, another means by which to gauge the engagement of a participant could be to introduce a device that measures physiological signals. For

example, using a sensor device to measure heart rate variability or galvanic skin response would facilitate the measurement of natural physiological reactions to the education. This would augment the types of engagement behaviours recorded and enable a multifaceted evaluation of user engagement to be undertaken.

- The association between health literacy and an individual's capability to obtain, understand and use health information has been a major theme of this PhD. The first study (Chapter 3) sought to appreciate the relationships between online health information seeking behaviours and an individual's health literacy skills. Accordingly, during the second study (Chapter 4) a personalised patient education system was developed which aimed to provide education at a readability level suitable for a person's health literacy needs. This system was further utilised in the study of patient engagement (Chapter 6). Thereby, measuring a participant's health literacy skills and including this in data analysis would be a complementary extension to the patient engagement study. Similar studies in the sphere of health literacy research have sought to identify whether there are associations between health literacy and approaches to reading, evaluating and processing health information [58, 461]. However, to date, a study similar to this, which compared engagement with generic and personalised patient education, has not been identified in this research area. Including health literacy measurement would add a further dimension to the patient engagement study by provisioning investigation of associations between health literacy and engagement with features of the generic and personalised education. Moreover, this would make a novel contribution to knowledge in health literacy research.

Collaborative Ontology Evaluation (Chapter 5)

- Chapter 5 described a study which aimed to include domain experts with variable modelling experience in the evaluation of the patient education ontology. However, there was variability amongst the participants in their level of engagement with all the engineering task types. One means by which to further assist the participants to undertake a wider range of tasks could be to provide more supporting technical information. This could include, for example, extending the screencast to illustrate how to complete the different tasks within the survey interface. The screencast currently uses an example of entities within a university to describe different

ontological concepts. However, although these concepts may be familiar to the participants, changing the example to focus on a common health condition may provide greater assistance to the participants in understanding how to structure clinical knowledge within an ontology.

7.6 Recommendations for Similar Research Studies

During this PhD four studies were completed which focused on different aspects of health information provision. As described in this thesis, several challenges were encountered during these studies. A number of recommendations are now proposed which aim to assist researchers that are undertaking similar studies.

- During this PhD three studies were completed that included human subjects. However, for all three studies, there was a lack of diversity in the age and educational background of the study population. Recruitment for these studies mostly took place on a university campus, and therefore the pool of potential subjects tended to be younger and well-educated. As described in Chapter 3 and Chapter 6, it is difficult to determine to extent to which these characteristics impacted upon the study results. Thereby, the first recommendation for carrying out similar research is to focus on recruiting participants with a suitably wide age range and increased scope of academic achievement. This will require that recruitment activities are not focused solely on a university campus, but would include population groups that typically have a wider range of educational and work experiences. For example, recruiting amongst groups such as mother and baby groups, adult education and training classes, or older adult activity programmes may provide more variation in terms of age, education, experience of using Internet applications and health experiences.
- Another consequence of recruiting on a university campus was that some of the participants were colleagues of the investigator. As detailed in Chapter 3 and Chapter 6, this may have added some stress to the participants or may have encouraged their enthusiasm to perform well in the experiments. Thereby, the second recommendation is to avoid recruiting subjects from within the investigator's immediate working environment. Although it is convenient to recruit participants who are familiar with undertaking experiments of this type, it was found

in this PhD, that having a high quantity of subjects with a technical background likely contributed to several confounding factors in the studies.

- The two studies described in Chapter 3 and Chapter 6 required the participants to attend an experiment at which the investigator was also present. Again, there is some uncertainty as to how the presence of the investigator may have influenced the behaviour of the subjects. Thereby, the third recommendation is to organise the environment in which the experiment takes place, to assist the participant to be less aware of the investigator's presence. One possible means to achieve this could be to have the investigator sit in an adjoining room during the experiment, whilst still being accessible to the participant if assistance is required. However, it must also be acknowledged that the ability to implement this recommendation would be reliant on the layout of the environment in which the experiment takes place.
- One of the main outputs from this PhD is an ontology-based architecture for personalised patient education. However, whilst the patient education ontology supported the production of personalised education, it was time-consuming to develop and, as described in Chapter 5, several issues were encountered during evaluation of the ontology. Thereby the fourth recommendation relates to the development of ontology-based architectures, and proposes that before deciding to design a novel ontology, the engineers should investigate whether a domain ontology could be utilised to achieve the required functionality in the architecture. This would reduce the time and effort used to develop a novel ontology, and may reduce concerns regarding the validity and correctness of the ontological model.

7.7 Recommendations for Health Information Provision

The Internet is a prominent source of both certified and unaccredited health information. This thesis has highlighted some of the strategies that information seekers employ when utilising online health information, and electronic and paper formats of patient education. Several recommendations are now proposed which relate to the provision and use of electronic formats of health information.

- Chapter 3 described a study of online health information seeking. The study presented subjects with questions related to popular health topics. The subjects predominantly relied on search engines to locate health resources, a common behaviour amongst online health information seekers [23]. Other prominent characteristics of the study population included a reliance on unaccredited health information, and a reluctance to confirm the validity of health information by comparing it with other sources. Although in many cases the unaccredited information was accurate, these behavioural traits are concerning within such a highly qualified and confident group of health information seekers. Online health information is accessed by individuals of all ages, and socio-economic and educational backgrounds [23], thereby not all information seekers will have skills comparable to these participants to interpret, critique and apply the information found. One possible approach to alleviating the problems associated with limited health information seeking skills could be public health promotion of criteria by which to assess the validity of online health information. Providing standardised evaluation criteria could assist information seekers of all abilities to effectively identify unreliable information sources. These criteria could be promoted through posters and leaflets in physician surgeries and health organisations, and on online platforms such as NHS Choices. There is also an opportunity to enhance the health literacy and eHealth literacy skills of adolescents and young adults through education in schools, colleges and universities. This also has the benefit of equipping younger generations with eHealth literacy skills that can inform their health decisions as their health needs change throughout adult life.
- Chapter 6 described a study which evaluated how users engage with generic and personalised forms of patient education. Analysis of the participants' gaze behaviour provided some insights as to which elements of education may be most effective in attracting and retaining the attention of readers. It was found, that regardless of the type of education being used, the subjects spent most time reading the text content. This emphasises that for many patients the text information is the primary focus, and therefore it is essential that text content is clear and understandable for the reader. The widespread use of online health resources provides an opportunity for health organisations to improve the usefulness of patient

education by adapting the content to be more engaging and comprehensible for individuals. For example, Internet based applications could provide health information that is transcribed as suitable for a patient's literacy level. The inclusion of personalised images that a reader can identify with may also encourage engagement with the education. Humans have a tendency to focus their attention on people within images and illustrations, and this provides an opportunity to draw the attention of readers to particular images. For example, including an image of a person carrying out a diabetic health check, such as a blood glucose check, may attract the reader's attention and reinforce the importance of this health check. The level of attention focused on the diagram (included in the generic booklet) during the study indicates that including cognitive illustrations could also be an effective method of drawing the interest of readers to important points of information. Moreover, utilising web-based technologies can also enable the provision of cognitive illustrations that are suitable to the health literacy level of the reader, which can further enhance the usefulness of the education.

7.8 Conclusion

This thesis presented a thorough investigation of the human factors associated with using health information, and the technological challenges associated with delivering health information. The research began by focusing on the growing use of the Internet as a source of health information, and the first research project addressed the human activities and skills involved with locating and using online health information. Subsequently, the focus of investigation shifted to the use of patient education materials, and it was considered whether a novel tailoring strategy could enhance the usability and attractiveness of the education for patients. In parallel, it was also evaluated whether semantic web technologies could be used to provide electronic education that was adaptive to the different personal characteristics, health status and literacy needs of patients. Chapter 4 described the design and implementation of an ontology based architecture which created personalised education for diabetic patients. Following this, Chapter 5 sought to strengthen the robustness of this electronic solution by verifying the validity of the ontological model. The research focus then progressed to consider user engagement with different formats of patient education. The study described in Chapter 6 aimed to evaluate whether the engagement behaviours displayed by individuals when using generic or personalised formats of patient education would

compare or contrast. Furthermore, the study also sought to determine whether an individual's level of engagement would fluctuate in response to viewing different personalised features in electronic education. The research objectives of this PhD project were fulfilled through these four studies and the completion of a review of related literature (as detailed in Chapter 2). Section 7.2 explicates precisely the manner in which these objectives were achieved. However, due to constraints of time and geography several limitations arose for the three studies that involved human subjects (Chapter 3, 5 and 6). During these studies recruitment activities were mainly targeted at university staff and students, and consequently there was reduced diversity within the study populations. However, for these three studies, the volume and range of the collected data set ensured that the results found were credible and robust. Reviewing the methodology and research outputs from these studies also identified various areas for future development as described in Section 7.5. Applying these modifications could expand the scope of the research, and augment and enrich the research findings.

This thesis aimed to *“use software tracking and user log analysis to appraise the health seeking behaviours and skills involved in utilising and engaging with web-based health information applications, and in turn to develop a computer-enhanced personalised patient education system using ontologies and related technologies”*. This aim has been achieved through the research completed throughout this PhD project, as recounted in this thesis. Comprehensible and accurate health information is a valuable resource for citizens as they manage concerns about their own health and navigate modern healthcare systems. Advances in electronic communications have enabled ubiquitous access to Internet-based health information services, and consumption of this information is ever increasing. Moreover, novel technologies, including semantic web technologies, have enabled the provision of electronic health tools that can adapt dynamically to fulfil the personal health information needs of consumers. Government health strategies worldwide have embraced the use of information technology as a means to support health information provision, whilst potentially improving patient experience and reducing the pressure on frontline services. However, this thesis has underlined the complexity of challenges that individuals encounter when attempting to engage with health information. Appreciating the personal characteristics and skills of citizens is a fundamental issue when considering the use of information technology to deliver health information services. Individuals will have unique abilities and health

experiences, thereby personalising information to focus on the particular needs of each person will enhance its usability and efficacy. Delivering information services of this type requires technology that can represent the complexity of human characteristics and health experiences. Semantic web technologies provide the sophistication to capture this range of data and enable tailoring services. Realising electronic health information applications that can fulfil the diverse personal and health needs of consumers will assist citizens in attaining the manifold benefits that are possible from using high quality health information.

References

- [1] C. Hoving, A. Visser, P.D. Mullen and B. van den Borne, "A history of patient education by health professionals in Europe and North America: From authority to shared decision making education," *Patient Educ. Couns.*, vol. 78, pp. 275-281, 3, 2010.
- [2] E.J. Emanuel and L.L. Emanuel, "Four models of the physician-patient relationship," *Jama*, vol. 267, pp. 2221-2226, 16, 1992.
- [3] Department of Health (2015). *The NHS Constitution for England*. [Online]. Available: <https://www.gov.uk/government/publications/the-nhs-constitution-for-england/the-nhs-constitution-for-england>. [Accessed: June 16, 2017].
- [4] Agency for Healthcare Research and Quality (2015, July). *Shared Decision-Making*. [Online]. Available: <https://www.ahrq.gov/cahps/quality-improvement/improvement-guide/6-strategies-for-improving/communication/strategy6i-shared-decisionmaking.html>. [Accessed: June 16, 2017].
- [5] NHS England, Care Quality Commission, Health Education England, Monitor, NHS Trust Development Authority and Public Health England, "NHS Five Year Forward View," NHS England, London, 2014. [Online]. Available: <https://www.england.nhs.uk/five-year-forward-view/>. [Accessed: June 16, 2017].
- [6] Committee on Quality of Health Care in America, Institute of Medicine, "Crossing the Quality Chasm: A New Health System for the 21st Century," National Academies Press, 2001 [Online]. Available: <https://www.nap.edu/catalog/10027/crossing-the-quality-chasm-a-new-health-system-for-the>. [Accessed: June 16, 2017].
- [7] World Health Organization, "Patient Engagement: Technical Series on Safer Primary Care," World Health Organization, Geneva, 2016. [Online]. Available: <http://apps.who.int/iris/bitstream/10665/252269/1/9789241511629-eng.pdf?ua=1>. [Accessed: June 16, 2017].
- [8] European Commission, "Together for Health: A Strategic Approach for the EU 2008-2013," European Commission, 2007. [Online]. Available: https://ec.europa.eu/health/ph_overview/Documents/strategy_wp_en.pdf. [Accessed: June 16, 2017].

- [9] Department of Health, "Your health, your way - A guide to long term conditions and self care. Information for healthcare professionals," Crown copyright, 2009. [Online]. Available: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_097588. [Accessed: June 16, 2017].
- [10] Department of Health, "Supporting people with long term conditions: commissioning personalised care planning – a guide for commissioners," Crown copyright, 2009. [Online]. Available: http://webarchive.nationalarchives.gov.uk/+//www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_093354. [Accessed: June 16, 2017].
- [11] A. Coulter and A. Collins, "Making shared decision-making a reality. No decision about me, without me," The King's Fund, London, 2011. [Online]. Available: <https://www.kingsfund.org.uk/publications/making-shared-decision-making-reality>. [Accessed: June 16, 2017].
- [12] NHS England (n.d.). *Shared Decision Making*. [Online]. Available: <https://www.england.nhs.uk/ourwork/pe/sdm/>. [Accessed: June 16, 2017].
- [13] Department of Health, "Help and advice for living life well with a long term condition," Crown copyright, 2009. [Online]. Available: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_102996. [Accessed: June 16, 2017].
- [14] B. Gaglio, R. E. Glasgow and S. S. Bull, "Do Patient Preferences for Health Information Vary by Health Literacy or Numeracy? A Qualitative Assessment," *J. Health Commun.*, vol. 17 Suppl. 3, pp. 109-121, 2012.
- [15] M.E. Mills and K. Sullivan, "The importance of information giving for patients newly diagnosed with cancer: a review of the literature," *J. Clin. Nurs.*, vol. 8, pp. 631-642, 6, 1999.
- [16] Department of Health, "Supporting people with long term conditions to self care – A guide to developing local strategies and good practice," Crown copyright, 2006. [Online]. Available: <http://webarchive.nationalarchives.gov.uk/+//dh.gov.uk/en/publicationsandstatistics>

- /publications/publicationspolicyandguidance/dh_4130725. [Accessed: June 16, 2017].
- [17] D. Stacey, F. Légaré, K. Lewis, M.J. Barry, C.L. Bennett, K. B. Eden, M. Holmes-Rovner, H. Llewellyn-Thomas, A. Lyddiatt, R. Thomson and L. Trevena, "Decision aids for people facing health treatment or screening decisions," *Cochrane Database of Systematic Reviews*, Issue 4, 2017. doi: 10.1002/14651858.CD001431.pub5.
 - [18] O. Husson, F. Mols and L.V. van de Poll-Franse, "The relation between information provision and health-related quality of life, anxiety and depression among cancer survivors: a systematic review," *Ann. Oncol.*, vol. 22, pp. 761-772, 4, 2011.
 - [19] B.W. Jack, V.K. Chetty, D. Anthony, J.L. Greenwald, G.M. Sanchez, A.E. Johnson, S.R. Forsythe, J.K. O'Donnell, M.K. Paasche-Orlow and C. Manasseh, "A reengineered hospital discharge program to decrease rehospitalization: a randomized trial," *Ann. Intern. Med.*, vol. 150, pp. 178-187, 3, 2009.
 - [20] A. Mulley, C. Trimble and G. Elwyn, "Patients' preferences matter. Stop the silent misdiagnosis," The King's Fund, London, 2012 [Online]. Available: <https://www.kingsfund.org.uk/publications/patients-preferences-matter>. [Accessed: June 16, 2017].
 - [21] P. M. Massey, "Where Do U.S. Adults Who Do Not Use the Internet Get Health Information? Examining Digital Health Information Disparities From 2008 to 2013," *J. Health Commun.*, vol. 21, pp. 118-124, 1, 2016.
 - [22] European Commission, "European citizens' digital health literacy," European Commission, 2014. [Online]. Available: http://ec.europa.eu/public_opinion/flash/fl_404_sum_en.pdf. [Accessed: June 16, 2017].
 - [23] S. Fox and M. Duggan, "Health Online," Pew Research Center, Washington, DC, 2013. [Online]. Available: <http://www.pewinternet.org/2013/01/15/health-online-2013/>. [Accessed: June 16, 2017].
 - [24] M. Pohjanoksa-Mäntylä, J.S. Bell, S. Helakorpi, U. Närhi, A. Pelkonen and M.S. Airaksinen, "Is the Internet replacing health professionals? A population survey on sources of medicines information among people with mental disorders," *Soc. Psychiatry Psychiatr. Epidemiol.*, vol. 46, pp. 373-379, 5, 2011.

- [25] Á. Mendes, L. Abreu, M.R. Vilar-Correia and J. Borlido-Santos, "“That Should Be Left to Doctors, That’s What They are There For!”—Exploring the Reflexivity and Trust of Young Adults When Seeking Health Information," *Health Commun.*, pp. 1-6, 2016. doi: <http://dx.doi.org/10.1080/10410236.2016.1199081>.
- [26] E. M. Mathew, K. Rajiah and K. K. Sharma, "Interpretation of consumer's perception on readability of Consumer Medical Information Leaflets on obesity and lipid lowering drugs with standard methods," *Journal of Pharmacy Research*, vol. 7, pp. 606-610, 7, 2013.
- [27] S. Speirs, S. Rees and M. Tagoe, "An audit of foot surgery information leaflets from the patients’ perspective," *The Foot*, vol. 18, pp. 7-14, 1, 2008.
- [28] S. Fox, "Health Topics," Pew Research Center, Washington, DC, 2011.
[Online]. Available: <http://www.pewinternet.org/2011/02/01/health-topics-2/>.
[Accessed: June 16, 2017].
- [29] A. Prestin, S.N. Vieux and W.S. Chou, "Is online health activity alive and well or flatlining? Findings from 10 years of the Health Information National Trends Survey," *J. Health Commun.*, vol. 20, pp. 790-798, 7, 2015.
- [30] K. Sørensen, S. Van den Broucke, J. Fullam, G. Doyle, J. Pelikan, Z. Slonska and H. Brand, "Health literacy and public health: A systematic review and integration of definitions and models," *BMC Public Health*, vol. 12, pp. 1-13, 2012.
- [31] L. Nielsen-Bohlman, A. M. Panzer, and D. A. Kindig (Eds.). "Health literacy: a prescription to end confusion," National Academies Press, 2004. [Online]. Available: <https://psnet.ahrq.gov/resources/resource/1147>. [Accessed: June 16, 2017].
- [32] D. Nutbeam, "Health Promotion Glossary," World Health Organization, Geneva, 1998. [Online]. Available: <http://www.who.int/healthpromotion/about/HPG/en/>. [Accessed: June 16, 2017].
- [33] D. Nutbeam, "The evolving concept of health literacy," *Soc. Sci. Med.*, vol. 67, pp. 2072-2078, 12, 2008.
- [34] K. Sørensen, J.M. Pelikan, F. Rothlin, K. Ganahl, Z. Slonska, G. Doyle, J. Fullam, B. Kondilis, D. Agrafiotis, E. Uiters, M. Falcon, M. Mensing, K. Tchamov, S. van den Broucke, H. Brand and HLS-EU Consortium, "Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU)," *Eur. J. Public Health*, vol. 25, pp. 1053-1058, 6, 2015.

- [35] M. Kutner, E. Greenberg, Y. Jin and C. Paulsen, "The Health Literacy of America's Adults: Results from the 2003 National Assessment of Adult Literacy (NCES 2006-483)," U.S. Department of Education. Washington, DC, 2006. [Online]. Available: <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006483>. [Accessed: June 16, 2017].
- [36] I. Kickbusch, J. M. Pelikan, F. Apfel and A. D. Tsouros, "Health literacy. The solid facts," World Health Organization, Geneva, 2013. [Online]. Available: <http://www.euro.who.int/en/publications/abstracts/health-literacy.-the-solid-facts>. [Accessed: June 16, 2017].
- [37] M. Wei, "The associations between health literacy, reasons for seeking health information, and information sources utilized by Taiwanese adults," *Health Educ. J.*, vol. 73, pp. 423-434, 4, 2014.
- [38] T. Olives, R. Patel, S. Patel, J. Hottinger and J.R. Miner, "Health literacy of adults presenting to an urban ED," *Am. J. Emerg. Med.*, vol. 29, pp. 875-882, 10, 2011.
- [39] C. Wang, R.L. Kane, D. Xu and Q. Meng, "Health literacy as a moderator of health-related quality of life responses to chronic disease among Chinese rural women," *BMC Womens Health*, vol. 15, pp. 34, 2015.
- [40] N. D. Berkman, S.L. Sheridan, K.E. Donahue, D.J. Halpern and K. Crotty, "Low health literacy and health outcomes: an updated systematic review," *Ann. Intern. Med.*, vol. 155, pp. 97-107, 2, 2011.
- [41] S. Bostock and A. Steptoe, "Association between low functional health literacy and mortality in older adults: longitudinal cohort study," *Bmj*, vol. 344, pp. e1602, 2012.
- [42] A. E. Schnitzer, M. Rosenzweig and B. Harris, "Health Literacy: A Survey of the Issues and Solutions," *Journal of Consumer Health on the Internet*, vol. 15, pp. 164-179, 2, 2011.
- [43] J. H. Fan, S.A. Lyons, M.S. Goodman, M.S. Blanchard and K.A. Kaphingst, "Relationship Between Health Literacy and Unintentional and Intentional Medication Nonadherence in Medically Underserved Patients With Type 2 Diabetes," *Diabetes Educ.*, vol. 42, pp. 199-208, 2, 2016.
- [44] S. E. Mitchell, E. Sadikova, B.W. Jack and M. Paasche-Orlow, "Health Literacy and 30-Day Postdischarge Hospital Utilization," *J. Health Commun.*, vol. 17 Suppl. 3, pp. 325-338, 2012.

- [45] Health Literacy Centre Europe (2015, Nov). *Costs of low health literacy*. [Online]. Available: <http://healthliteracycentre.eu/costs-of-low-health-literacy/>. [Accessed: June 16, 2017].
- [46] D.W. Baker, J.A. Gazmararian, M.V. Williams, T. Scott, R.M. Parker, D. Green, J. Ren and J. Peel, "Functional health literacy and the risk of hospital admission among Medicare managed care enrollees," *Am. J. Public Health*, vol. 92, pp. 1278-1283, 8, 2002.
- [47] Health Literacy Centre Europe (2015, Nov). *National Health Literacy policies and what we can learn from them*. [Online]. Available: <http://healthliteracycentre.eu/national-health-literacy-policies/>. [Accessed: June 16, 2017].
- [48] G. Quaglio, K. Sorensen, P. Rubig, L. Bertinato, H. Brand, T. Karapiperis, I. Dinca, T. Peetso, K. Kadenbach and C. Dario, "Accelerating the health literacy agenda in Europe," *Health. Promot. Int.*, 2016. doi: <https://doi.org/10.1093/heapro/daw028>
- [49] Office of Disease Prevention and Health Promotion (2017). *Health Communication and Health Information Technology*. [Online]. Available: <https://www.healthypeople.gov/2020/topics-objectives/topic/health-communication-and-health-information-technology> [Accessed: June 16, 2017].
- [50] U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, "National Action Plan to Improve Health Literacy," U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, Washington, DC, 2010. [Online]. Available: <https://www.cdc.gov/healthliteracy/planact/national.html>. [Accessed: June 16, 2017].
- [51] European Commission (n.d.). *Health programme*. [Online]. Available: https://ec.europa.eu/health/programme/policy_en. [Accessed: June 16, 2017].
- [52] A. M. H. Chen, K.S. Yehle, N.M. Albert, K.F. Ferraro, H.L. Mason, M.M. Murawski and K.S. Plake, "Relationships between health literacy and heart failure knowledge, self-efficacy, and self-care adherence," *Research in Social and Administrative Pharmacy*, vol. 10, pp. 378-386, 2, 2014.
- [53] C. Y. Osborn, M.K. Paasche-Orlow, S.C. Bailey and M.S. Wolf, "The mechanisms linking health literacy to behavior and health status," *Am. J. Health Behav.*, vol. 35, pp. 118-128, 1, 2011.

- [54] N. S. Morris, T.S. Field, J.L. Wagner, S.L. Cutrona, D.W. Roblin, B. Gaglio, A.E. Williams, P.J.K. Han, M.E. Costanza and K.M. Mazor, "The Association Between Health Literacy and Cancer-Related Attitudes, Behaviors, and Knowledge," *J. Health Commun.*, vol. 18 Suppl. 1, pp. 223-241, 2013.
- [55] J.A. Gazmararian, M.V. Williams, J. Peel and D.W. Baker, "Health literacy and knowledge of chronic disease," *Patient Educ. Couns.*, vol. 51, pp. 267-275, 3, 2003.
- [56] H. S. Yin, M. Johnson, A.L. Mendelsohn, M.A. Abrams, L.M. Sanders and B.P. Dreyer, "The health literacy of parents in the United States: a nationally representative study," *Pediatrics*, vol. 124 Suppl. 3, pp. S289-98, 2009.
- [57] C. von Wagner, C. Semmler, A. Good and J. Wardle, "Health literacy and self-efficacy for participating in colorectal cancer screening: The role of information processing," *Patient Educ. Couns.*, vol. 75, pp. 352-357, 6, 2009.
- [58] C. S. Meppelink and N. Bol, "Exploring the role of health literacy on attention to and recall of text-illustrated health information: An eye-tracking study," *Comput. Hum. Behav.*, vol. 48, pp. 87-93, 7, 2015.
- [59] S. D. Persell, C.Y. Osborn, R. Richard, S. Skripkauskas and M.S. Wolf, "Limited Health Literacy is a Barrier to Medication Reconciliation in Ambulatory Care," *Journal of General Internal Medicine*, vol. 22, pp. 1523-1526, 11, 2007.
- [60] S. E. Lillie, N.T. Brewer, S.C. O'Neill, E.F. Morrill, E.C. Dees, L.A. Carey and B.K. Rimer, "Retention and Use of Breast Cancer Recurrence Risk Information from Genomic Tests: The Role of Health Literacy," *Cancer Epidemiology Biomarkers & Prevention*, vol. 16, pp. 249-255, 2, 2007.
- [61] D.M. McCarthy, K.R. Waite, L.M. Curtis, K.G. Engel, D.W. Baker and M.S. Wolf, "What did the doctor say? Health literacy and recall of medical instructions," *Med. Care*, vol. 50, pp. 277-282, 4, 2012.
- [62] S. M. Halbach, N. Ernstmann, C. Kowalski, H. Pfaff, T. Pförtner, S. Wesselmann and A. Enders, "Unmet information needs and limited health literacy in newly diagnosed breast cancer patients over the course of cancer treatment," *Patient Educ. Couns.*, vol. 99, pp. 1511-1518, 9, 2016.
- [63] M. Heijmans, G. Waverijn, J. Rademakers, R. van der Vaart and M. Rijken, "Functional, communicative and critical health literacy of chronic disease patients and their importance for self-management," *Patient Educ. Couns.*, vol. 98, pp. 41-48, 1, 2015.

- [64] M. Paasche-Orlow, K.A. Riekert, A. Bilderback, A. Chanmugam, P. Hill, C.S. Rand, F.L. Brancati and J.A. Krishnan, "Tailored Education May Reduce Health Literacy Disparities in Asthma Self-Management," *Am. J. Respir. Crit. Care Med.*, vol. 172, pp. 980-986, 8, 2005.
- [65] A. Shaw, S. Ibrahim, F. Reid, M. Ussher and G. Rowlands, "Patients' perspectives of the doctor–patient relationship and information giving across a range of literacy levels," *Patient Educ. Couns.*, vol. 75, pp. 114-120, 4, 2009.
- [66] Department of Health (2015). *NHS choices. Your health, your choices*. [Online]. Available: <http://www.nhs.uk/pages/home.aspx>. [Accessed: June 16, 2017].
- [67] Paul Houghton (2016, June). *NHS Choices Blog. Choices scale blog. Data, Home, Technical*. [Online]. Available: <http://blogs.nhs.uk/choices-blog/2016/06/17/choices-scale-blog/>. [Accessed: June 16, 2017].
- [68] U.S. National Library of Medicine (2017, June). *MedlinePlus. Trusted Health Information for You*. [Online]. Available: <https://medlineplus.gov/>. [Accessed: June 16, 2017].
- [69] European Commission, "eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century," European Commission, 2012. [Online]. Available: https://ec.europa.eu/health/sites/health/files/ehealth/docs/com_2012_736_en.pdf. [Accessed: January 10, 2018].
- [70] World Health Organization, "Global diffusion of eHealth: Making universal health coverage achievable. Report of the third global survey on eHealth," World Health Organization, Geneva, 2016. [Online]. Available: http://www.who.int/goe/publications/global_diffusion/en/ [Accessed: June 16, 2017].
- [71] National Information Board and Department of Health, "Personalised Health and Care 2020: a framework for action," National Information Board and Department of Health, 2014 [Online]. Available: <https://www.gov.uk/government/publications/personalised-health-and-care-2020>. [Accessed: June 16, 2017].
- [72] Tinder Foundation, "Health & Digital: Reducing Inequalities, Improving Society. An evaluation of the Widening Digital Participation programme," Tinder Foundation, 2016. [Online]. Available: <http://nhs.tinderfoundation.org/wp->

- content/uploads/2016/07/Improving_Digital_Health_Skills_Report_2016.pdf.
[Accessed: June 16, 2017].
- [73] A. Dobrev, T. Jones, K. Stroetmann, Y. Vatter and K. Peng, "The socio-economic impact of interoperable electronic health record (EHR) and ePrescribing systems in Europe and beyond," European Commission, 2009. [Online]. Available: <https://ec.europa.eu/digital-single-market/en/news/socio-economic-impact-interoperable-electronic-health-record-ehr-and-eprescribing-systems>. [Accessed: June 16, 2017].
- [74] M. Honeyman, P. Dunn and H. McKenna, "A digital NHS? An introduction to the digital agenda and plans for implementation," The King's Fund, London, 2016. [Online]. Available: <https://www.kingsfund.org.uk/publications/digital-nhs>. [Accessed: June 16, 2017].
- [75] T. W. Bickmore and M. K. Paasche-Orlow, "The role of information technology in health literacy research," *J. Health Commun.*, vol. 17 Suppl. 3, pp. 23-29, 2012.
- [76] European Commission, "eHealth Task Force Report – Redesigning health in Europe for 2020," Publications Office of the European Union, Luxembourg, 2012 [Online]. Available: <http://www.eu-patient.eu/News/News-Archive/Report-Redesigning-health-in-Europe-for-2020/>. [Accessed: June 16, 2017]
- [77] S. Quinn (2016). *hciParser*. [Software]. Available: <https://github.com/squinn36/hciParser> [Accessed: June 16, 2017].
- [78] F. Smith, E. Carlsson, D. Kokkinakis, M. Forsberg, K. Kodeda, R. Sawatzky, F. Friberg and J. Öhlén, "Readability, suitability and comprehensibility in patient education materials for Swedish patients with colorectal cancer undergoing elective surgery: A mixed method design," *Patient Educ. Couns.*, vol. 94, pp. 202-209, 2, 2014.
- [79] D. Hirsh, R. Clerehan, M. Staples, R.H. Osborne and R. Buchbinder, "Patient assessment of medication information leaflets and validation of the Evaluative Linguistic Framework (ELF)," *Patient Educ. Couns.*, vol. 77, pp. 248-254, 11, 2009.
- [80] N. Cox, C. Bowmer and A. Ring, "Health literacy and the provision of information to women with breast cancer," *Clin. Oncol.*, vol. 23, pp. 223-227, 3, 2011.

- [81] A. Cawsey, F. Grasso and C. Paris, "Adaptive Information for Consumers of Healthcare," in *The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds. Berlin: Springer, 2007, pp. 465-484.
- [82] T. C. Davis, M.S. Wolf, P.F. Bass, J.A. Thompson, H.H. Tilson, M. Neuberger and R.M. Parker, "Literacy and Misunderstanding Prescription Drug Labels," *Ann. Intern. Med.*, vol. 145, pp. 887-894, 12, 2006.
- [83] H. Pander Maat and L. Lentz, "Improving the usability of patient information leaflets," *Patient Educ. Couns.*, vol. 80, pp. 113-119, 7, 2010.
- [84] B. K. Kondilis, P.D. Akrivos, T. Sardi, E. Soteriades and M. Falagas, "Readability levels of health pamphlets distributed in hospitals and health centres in Athens, Greece," *Public Health*, vol. 124, pp. 547-552, 10, 2010.
- [85] J. K. Vallance, L. M. Taylor and C. Lavallee, "Suitability and readability assessment of educational print resources related to physical activity: implications and recommendations for practice," *Patient Educ. Couns.*, vol. 72, pp. 342-349, 2, 2008.
- [86] K. Joubert and E. Githinji, "Quality and readability of information pamphlets on hearing and paediatric hearing loss in the Gauteng Province, South Africa," *Int. J. Pediatr. Otorhinolaryngol.*, vol. 78, pp. 354-358, 2, 2014.
- [87] C. Haw and J. Stubbs, "Patient information leaflets for antidepressants: Are patients getting the information they need?" *J. Affect. Disord.*, vol. 128, pp. 165-170, 1-2, 2011.
- [88] J. Powell, N. Inglis, J. Ronnie and S. Large, "The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study," *J. Med. Internet Res.*, vol. 13, pp. e20, 1, 2011.
- [89] A. Bianco, R. Zucco, C.G.A. Nobile, C. Pileggi and M. Pavia, "Parents seeking health-related information on the Internet: cross-sectional study," *J. Med. Internet Res.*, vol. 15, pp. e204, 9, 2013.
- [90] M. Lemire, G. Paré, C. Sicotte and C. Harvey, "Determinants of Internet use as a preferred source of information on personal health," *Int. J. Med. Inform.*, vol. 77, pp. 723-734, 11, 2008.
- [91] S. Fox and M. Duggan, "The Diagnosis Difference," Pew Research Center, Washington, DC, 2013. [Online]. Available:

- <http://www.pewinternet.org/2013/11/26/the-diagnosis-difference/> [Accessed: June 16, 2017].
- [92] M. Caiata-Zufferey, A. Abraham, K. Sommerhalder and P.J. Schulz, "Online Health Information Seeking in the Context of the Medical Consultation in Switzerland," *Qual. Health Res.*, vol. 20, pp. 1050-1061, 8, 2010.
 - [93] F. A. Stevenson, C. Kerr, E. Murray and I. Nazareth, "Information from the Internet and the doctor-patient relationship: the patient perspective – a qualitative study," *BMC Fam Pract.*, vol. 8, pp. 1-8, 2007.
 - [94] Y. Y. Yan, "Online Health Information Seeking Behavior in Hong Kong: An Exploratory Study," *J. Med. Syst.*, vol. 34, pp. 147-153, 2, 2010.
 - [95] K. Lee, K. Hoti, J.D. Hughes and L. Emmerton, "Consumer Use of “Dr Google”: A Survey on Health Information-Seeking Behaviors and Navigational Needs," *J. Med. Internet Res.*, vol. 17, pp. e288, 12, 2015.
 - [96] Y. A. Strekalova, "Finding Motivation: Online Information Seeking Following Newborn Screening for Cystic Fibrosis," *Qual. Health Res.*, vol. 26, pp. 1180-1190, 9, 2016.
 - [97] P. Sayakhot and M. Carolan-Olah, "Internet use by pregnant women seeking pregnancy-related information: a systematic review," *BMC Pregnancy and Childbirth*, vol. 16, pp. 1-10, 2016.
 - [98] P. Åsbring and A. L. Närvänen, "Patient power and control: a study of women with uncertain illness trajectories," *Qual. Health Res.*, vol. 14, pp. 226-240, 2, 2004.
 - [99] M. Lemire, C. Sicotte and G. Paré, "Internet use and the logics of personal empowerment in health," *Health Policy*, vol. 88, pp. 130-140, 10, 2008.
 - [100] M. Sajid, M. Iftikhar, R. Monteiro, A. Miles, W. Woods and M. Baig, "Internet information on colorectal cancer: commercialization and lack of quality control," *Colorectal Disease*, vol. 10, pp. 352-356, 4, 2008.
 - [101] M. Chung, R.P. Oden, B.L. Joyner, A. Sims and R.Y. Moon, "Safe Infant Sleep Recommendations on the Internet: Let's Google It," *J. Pediatr.*, vol. 161, pp. 1080-1084.e1, 6, 2012.
 - [102] P. A. Hendrick, O.H. Ahmed, S.S. Bankier, T.J. Chan, S.A. Crawford, C.R. Ryder, L.J. Welsh and A.G. Schneiders, "Acute low back pain information online: An evaluation of quality, content accuracy and readability of related websites," *Man. Ther.*, vol. 17, pp. 318-324, 8, 2012.

- [103] W. Chen and K. Lee, "More than search? Informational and participatory eHealth behaviors," *Comput. Hum. Behav.*, vol. 30, pp. 103-109, 1, 2014.
- [104] J. B. Weaver III, N.J. Thompson, S.S. Weaver and G.L. Hopkins, "Healthcare non-adherence decisions and internet health information," *Comput. Hum. Behav.*, vol. 25, pp. 1373-1380, 11, 2009.
- [105] J. Chae, "Online cancer information seeking increases cancer worry," *Comput. Hum. Behav.*, vol. 52, pp. 144-150, 11, 2015.
- [106] S. E. Baumgartner and T. Hartmann, "The Role of Health Anxiety in Online Health Information Search," *Cyberpsychology, Behavior, and Social Networking*, vol. 14, pp. 613-618, 10, 2011.
- [107] H. Liang, Y. Xue and S. K. Chase, "Online health information seeking by people with physical disabilities due to neurological conditions," *Int. J. Med. Inform.*, vol. 80, pp. 745-753, 11, 2011.
- [108] S. J. Chang and E. Im, "A path analysis of Internet health information seeking behaviors among older adults," *Geriatr. Nurs.*, vol. 35, pp. 137-141, 2, 2014.
- [109] E. Manafò and S. Wong, "Assessing the eHealth Literacy Skills of Older Adults: A Preliminary Study," *Journal of Consumer Health on the Internet*, vol. 16, pp. 369-381, 4, 2012.
- [110] C. Escoffery, K.R. Miner, D.D. Adame, S. Butler, L. McCormick and E. Mendell, "Internet use for health information among college students," *J. Am. Coll. Health.*, vol. 53, pp. 183-188, 4, 2005.
- [111] P. J. Schmitt and C. J. Prestigiacomo, "Readability of Neurosurgery-Related Patient Education Materials Provided by the American Association of Neurological Surgeons and the National Library of Medicine and National Institutes of Health," *World Neurosurgery*, vol. 80, pp. e33-e39, 5, 2013.
- [112] M. Arsenault, M. J. Blouin and M. J. Guitton, "Information quality and dynamics of patients' interactions on tonsillectomy web resources," *Internet Interventions*, vol. 4 Part 2, pp. 99-104, 2016.
- [113] S. Sabharwal, S. Badarudeen and S. Unes Kunju, "Readability of Online Patient Education Materials From the AAOS Web Site," *Clin. Orthop.*, vol. 466, pp. 1245-1250, 5, 2008.
- [114] P. F. Svider, N. Agarwal, O.J. Choudhry, A.F. Hajart, S. Baredes, J.K. Liu and J.A. Eloy, "Readability assessment of online patient education materials from

- academic otolaryngology–head and neck surgery departments," *Am. J. Otolaryngol.*, vol. 34, pp. 31-35, 1, 2013.
- [115] Y. Zhang, Y. Sun and B. Xie, "Quality of health information for consumers on the web: A systematic review of indicators, criteria, tools, and evaluation results," *J. Assn. Inf. Sci. Technol.*, vol. 66, pp. 2071-2084, 10, 2015.
- [116] E. Sillence, P. Briggs, P.R. Harris and L. Fishwick, "How do patients evaluate and make use of online health information?" *Soc. Sci. Med.*, vol. 64, pp. 1853-1862, 5, 2007.
- [117] G. Eysenbach and C. Kohler, "How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews," *Bmj*, vol. 324, pp. 573-577, 7337, 2002.
- [118] S. A. Rains and C. D. Karmikel, "Health information-seeking and perceptions of website credibility: Examining Web-use orientation, message characteristics, and structural features of websites," *Comput. Hum. Behav.*, vol. 25, pp. 544-553, 3, 2009.
- [119] T. Hong, "The influence of structural and message features on Web site credibility," *J. Am. Soc. Inf. Sci. Technol.*, vol. 57, pp. 114-127, 1, 2006.
- [120] L. Rozmovits and S. Ziebland, "What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs," *Patient Educ. Couns.*, vol. 53, pp. 57-64, 1, 2004.
- [121] A. E. Anker, A. M. Reinhart and T. H. Feeley, "Health information seeking: A review of measures and methods," *Patient Educ. Couns.*, vol. 82, pp. 346-354, 3, 2011.
- [122] S. D. Lambert and C. G. Loiselle, "Health information seeking behavior," *Qual. Health Res.*, vol. 17, pp. 1006-1019, 8, 2007.
- [123] E. N. Efthimiadis, "How students search for consumer health information on the web," in *2009 42nd Hawaii International Conference on System Sciences*, 2009, pp. 1-8.
- [124] E. G. Toms and C. Latter, "How consumers search for health information," *Health. Informatics J.*, vol. 13, pp. 223-235, 3, 2007.
- [125] M. Huang, D. Hansen and B. Xie, "Older adults' online health information seeking behavior," in *Proceedings of the 2012 iConference*, 2012, pp. 338-345.

- [126] D. L. Hansen, H.A. Derry, P.J. Resnick and C.R. Richardson, "Adolescents searching for health information on the Internet: an observational study," *J. Med. Internet Res.*, vol. 5, pp. e25, 4, 2003.
- [127] J. Sharit, J. Taha, R.W. Berkowsky, H. Profita and S.J. Czaja, "Online information search performance and search strategies in a health problem-solving scenario," *Journal of Cognitive Engineering and Decision Making*, vol. 9, pp. 211-228, 3, 2015.
- [128] Q. T. Zeng, S. Kogan, R.M. Plovnick, J. Crowell, E. Lacroix and R.A. Greenes, "Positive attitudes and failed queries: an exploration of the conundrums of consumer health information retrieval," *Int. J. Med. Inform.*, vol. 73, pp. 45-55, 2, 2004.
- [129] A. T. McCray and T. Tse, "Understanding search failures in consumer health information systems," in *Amia*, 2003, pp. 430-434.
- [130] R. van der Vaart, C.H. Drossaert, M. de Heus, E. Taal and M. van de Laar, "Measuring actual eHealth literacy among patients with rheumatic diseases: a qualitative analysis of problems encountered using Health 1.0 and Health 2.0 applications," *J. Med. Internet Res.*, vol. 15, pp. e27, 2, 2013.
- [131] C. C. Yang, F. Winston, M.A. Zarro and N. Kassam-Adams, "A study of user queries leading to a health information website: AfterTheInjury. org," in *Proceedings of the 2011 iConference*, 2011, pp. 267-272.
- [132] Q. Zeng, S. Kogan, N. Ash, R. Greenes and A. Boxwala, "Characteristics of consumer terminology for health information retrieval," *Methods Inf. Med.*, vol. 41, pp. 289-298, 4, 2002.
- [133] R. W. White and E. Horvitz, "Studies of the onset and persistence of medical concerns in search logs," in *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2012, pp. 265-274.
- [134] A. Spink, Y. Yang, J. Jansen, P. Nykanen, D.P. Lorence, S. Ozmutlu and H.C. Ozmutlu, "A study of medical and health queries to web search engines," *Health Info. Libr. J.*, vol. 21, pp. 44-51, 1, 2004.
- [135] Y. Zhang, P. Wang, A. Heaton and H. Winkler, "Health information searching behavior in MedlinePlus and the impact of tasks," in *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*, 2012, pp. 641-650.

- [136] Y. Zhang, "Searching for specific health-related information in MedlinePlus: Behavioral patterns and user experience," *J. Assn. Inf. Sci. Technol.*, vol. 65, pp. 53-68, 1, 2014.
- [137] A. Keselman, A. C. Browne and D. R. Kaufman, "Consumer Health Information Seeking as Hypothesis Testing," *J. Am. Soc. Inf. Sci. Technol.*, vol. 15, pp. 484-495, 4, 2008.
- [138] Q. T. Zeng and T. Tse, "Exploring and Developing Consumer Health Vocabularies," *Journal of the American Medical Informatics Association*, vol. 13, pp. 24-29, 1, 2006.
- [139] J. Zhang, D. Wolfram, P. Wang, Y. Hong and R. Gillis, "Visualization of health-subject analysis based on query term co-occurrences," *J. Am. Soc. Inf. Sci. Technol.*, vol. 59, pp. 1933-1947, 12, 2008.
- [140] M. S. Birru, V.M. Monaco, L. Charles, H. Drew, V. Njie, T. Bierria, E. Detlefsen and R.A. Steinman, "Internet usage by low-literacy adults seeking health information: an observational analysis," *J. Med. Internet Res.*, vol. 6, pp. e25, 3, 2004.
- [141] E. Sillence, P. Briggs, L. Fishwick and P. Harris, "Trust and mistrust of online health sites," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2004, pp. 663-670.
- [142] L. Graham, T. Tse and A. Keselman, "Exploring user navigation during online health information seeking," in *Amia*, 2006, pp. 299-303.
- [143] Y. Zhang, "Consumer health information searching process in real life settings," *Proc. Am. Soc. Info. Sci. Tech.*, vol. 49, pp. 1-10, 1, 2012.
- [144] D. Nutbeam, "Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century," *Health. Promot. Int.*, vol. 15, pp. 259-267, 3, 2000.
- [145] A. Peerson and M. Saunders, "Health literacy revisited: what do we mean and why does it matter?" *Health. Promot. Int.*, vol. 24, pp. 285-296, 3, 2009.
- [146] M. N. Verkissen, N.P.M. Ezendam, M.P. Fransen, M. Essink-Bot, M.J. Aarts, K.A.H. Nicolaije, M.C. Vos and O. Husson, "The role of health literacy in perceived information provision and satisfaction among women with ovarian tumors: A study from the population-based PROFILES registry," *Patient Educ. Couns.*, vol. 95, pp. 421-428, 6, 2014.

- [147] Y. A. Strekalova, "Seekers and avoiders: Using health information orientation to explore audience segmentation," *Journal of Communication in Healthcare*, vol. 7, pp. 228-237, 3, 2014.
- [148] J. M. Liechty, J. A. Saltzman and S. M. Musaad, "Health literacy and parent attitudes about weight control for children," *Appetite*, vol. 91, pp. 200-208, 2015.
- [149] M. Koo, I. Krass and P. Aslani, "Enhancing patient education about medicines: factors influencing reading and seeking of written medicine information," *Health Expectations*, vol. 9, pp. 174-187, 2, 2006.
- [150] J. R. Marks, J.M. Schectman, H. Groninger and M.L. Plews-Ogan, "The association of health literacy and socio-demographic factors with medication knowledge," *Patient Educ. Couns.*, vol. 78, pp. 372-376, 3, 2010.
- [151] N. Gutierrez, T.B. Kindratt, P. Pagels, B. Foster and N.E. Gimpel, "Health Literacy, Health Information Seeking Behaviors and Internet Use Among Patients Attending a Private and Public Clinic in the Same Geographic Area," *J. Community Health*, vol. 39, pp. 83-89, 1, 2014.
- [152] U. Sarkar, A.J. Karter, J.Y. Liu, N.E. Adler, R. Nguyen, A. Lopez and D. Schillinger, "The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system—results from the Diabetes Study of Northern California (DISTANCE)," *J. Health Commun.*, vol. 15 Suppl. 2, pp. 183-196, 2010.
- [153] J. Zoellner, C. Connell, W. Bounds, L. Crook and K. Yadrick, "Nutrition Literacy Status and Preferred Nutrition Communication Channels Among Adults in the Lower Mississippi Delta," *Preventing Chronic Disease*, vol. 6, pp. A128, 4, 2009.
- [154] N. Diviani, B. van den Putte, C.S. Meppelink and J.C.M. van Weert, "Exploring the role of health literacy in the evaluation of online health information: Insights from a mixed-methods study," *Patient Educ. Couns.*, vol. 99, pp. 1017-1025, 6, 2016.
- [155] R. J. Jacobs, J.Q. Lou, R.L. Ownby and J. Caballero, "A systematic review of eHealth interventions to improve health literacy," *Health. Informatics J.*, vol. 22, pp. 81-98, 2, 2016.
- [156] N. R. Kandula, P.A. Nsiah-Kumi, G. Makoul, J. Sager, C.P. Zei, S. Glass, Q. Stephens and D.W. Baker, "The relationship between health literacy and

- knowledge improvement after a multimedia type 2 diabetes education program," *Patient Educ. Couns.*, vol. 75, pp. 321-327, 6, 2009.
- [157] K. J. Silk, M.A. Horodyski, M. Rienzo, L. Mercer, B. Olson and R. Aldrich, "Strategies to increase health literacy in The Infant Feeding Series (TIFS): A Six-lesson curriculum for low-income mothers," *Health Promot. Pract.*, vol. 11, pp. 226-234, 2, 2010.
- [158] B. S. Gerber, I.G. Brodsky, K.A. Lawless, L.I. Smolin, A.M. Arozullah, E.V. Smith, M.L. Berbaum, P.S. Heckerling and A.R. Eiser, "Implementation and evaluation of a low-literacy diabetes education computer multimedia application," *Diabetes Care*, vol. 28, pp. 1574-1580, 7, 2005.
- [159] M. D. Bryant, E.D. Schoenberg, T.V. Johnson, M. Goodman, A. Owen-Smith and V.A. Master, "Multimedia version of a standard medical questionnaire improves patient understanding across all literacy levels," *J. Urol.*, vol. 182, pp. 1120-1125, 3, 2009.
- [160] T. W. Bickmore, L. M. Pfeifer and B. W. Jack, "Taking the time to care: Empowering low health literacy hospital patients with virtual nurse agents," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2009, pp. 1265-1274.
- [161] NHS England (n.d.). *The Information Standard*. [Online]. Available: <https://www.england.nhs.uk/tis/>. [Accessed: June 16, 2017].
- [162] Health On the Net Foundation (2016, June). *Health On the Net Foundation*. [Online]. Available: <https://www.hon.ch/>. [Accessed: June 16, 2017].
- [163] World Health Organization (2017). *Chronic diseases and health promotion*. [Online]. Available: http://www.who.int/chp/about/integrated_cd/en/. [Accessed: June 16, 2017].
- [164] National Center for Chronic Disease Prevention and Health Promotion (2016). *Chronic Disease Overview*. [Online]. Available: <https://www.cdc.gov/chronicdisease/overview/>. [Accessed: June 16, 2017].
- [165] R. P. Hawkins, M. Kreuter, K. Resnicow, M. Fishbein and A. Dijkstra, "Understanding tailoring in communicating about health," *Health Educ. Res.*, vol. 23, pp. 454-466, 3, 2008.
- [166] B. K. Rimer and M. W. Kreuter, "Advancing tailored health communication: A persuasion and message effects perspective," *J. Commun.*, vol. 56, S184-S201, 2006. doi: 10.1111/j.1460-2466.2006.00289.x

- [167] M. L. A. Lustria, J. Cortese, S.M. Noar and R.L. Glueckauf, "Computer-tailored health interventions delivered over the Web: review and analysis of key components," *Patient Educ. Couns.*, vol. 74, pp. 156-173, 2, 2009.
- [168] A. Jerant, N. Sohler, K. Fiscella, B. Franks and P. Franks, "Tailored interactive multimedia computer programs to reduce health disparities: opportunities and challenges," *Patient Educ. Couns.*, vol. 85, pp. 323-330, 2, 2011.
- [169] P. Krebs, J. O. Prochaska and J. S. Rossi, "A meta-analysis of computer-tailored interventions for health behavior change," *Prev. Med.*, vol. 51, pp. 214-221, 3-4, 2010.
- [170] S. M. Noar, C. N. Benac and M. S. Harris, "Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions." *Psychol. Bull.*, vol. 133, pp. 673, 4, 2007.
- [171] S. M. Noar, N. Grant Harrington, S.K. Van Stee and R. Shemanski Aldrich, "Tailored health communication to change lifestyle behaviors," *Am. J. Lifestyle Med.*, vol. 5, pp. 112-122, 2, 2011.
- [172] A. Oenema, J. Brug, A. Dijkstra, I. de Weerd and H. de Vries, "Efficacy and use of an internet-delivered computer-tailored lifestyle intervention, targeting saturated fat intake, physical activity and smoking cessation: a randomized controlled trial," *Ann. Behav. Med.*, vol. 35, pp. 125-135, 2, 2008.
- [173] A. Dijkstra, "The psychology of tailoring-ingredients in computer-tailored persuasion," *Social and Personality Psychology Compass*, vol. 2, pp. 765-784, 2, 2008.
- [174] R. A. Ruiter, L.T. Kessels, B.M. Jansma and J. Brug, "Increased attention for computer-tailored health communications: an event-related potential study." *Health Psychology*, vol. 25, pp. 300, 3, 2006.
- [175] A. Dijkstra, "Working mechanisms of computer-tailored health education: evidence from smoking cessation," *Health Educ. Res.*, vol. 20, pp. 527-539, 5, 2005.
- [176] M. Svensson and Y. T. Lagerros, "Motivational technologies to promote weight loss—From Internet to gadgets," *Patient Educ. Couns.*, vol. 79, pp. 356-360, 6, 2010.
- [177] B. Gislason, C. Mcknight, B. Potvin, S. Stuart, J. Zepeda, J. Weber and H. Elmiligi, "Introducing GlucoFit: An assistive technology for monitoring and managing diabetes," in *Proceedings of the 2012 Seventh International Conference*

- on Broadband, Wireless Computing, Communication and Applications*, 2012, pp. 414-419.
- [178] S. Hammer, J. Kim and E. André, "MED-StyleR: METABO diabetes-lifestyle recommender," in *Proceedings of the Fourth ACM Conference on Recommender Systems*, 2010, pp. 285-288.
 - [179] E. Chiauzzi, T.C. Green, S. Lord, C. Thum and M. Goldstein, "My student body: a high-risk drinking prevention web site for college students," *J. Am. Coll. Health.*, vol. 53, pp. 263-274, 6, 2005.
 - [180] L. Razmerita, T. Nabeth and K. Kirchner, "User modeling and attention support: Towards a framework of personalization techniques," in *CENTRIC 2012, the Fifth International Conference on Advances in Human-Oriented and Personalized Mechanisms, Technologies, and Services*, 2012, pp. 27-33.
 - [181] V. Wolowski, N. Ishikawa and H. Sumino, "Semantic web approach to content personalization," in *UBICOMM '07*, 2007, pp. 109-117.
 - [182] S. Islam, G. Freytag and R. Shankar, "Leveraging semantic web to retrieve customized medical information," in *Systems Conference (SysCon), 2013 IEEE International*, 2013, pp. 258-261.
 - [183] H. Bang and B. W. Wojdyski, "Tracking users' visual attention and responses to personalized advertising based on task cognitive demand," *Comput. Hum. Behav.*, vol. 55, pp. 867-876, 2016.
 - [184] F. Alt, A.S. Shirazi, A. Schmidt and J. Mennenöh, "Increasing the user's attention on the web: Using implicit interaction based on gaze behavior to tailor content," in *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense through Design*, 2012, pp. 544-553.
 - [185] L. T. Kessels, R.A. Ruiter, J. Brug and B.M. Jansma, "The effects of tailored and threatening nutrition information on message attention. Evidence from an event-related potential study," *Appetite*, vol. 56, pp. 32-38, 1, 2011.
 - [186] B. G. Danaher, S.M. Boles, L. Akers, J.S. Gordon and H.H. Severson, "Defining participant exposure measures in Web-based health behavior change programs," *J. Med. Internet Res.*, vol. 8, pp. e15, 3, 2006.
 - [187] V. J. Strecher, J. McClure, G. Alexander, B. Chakraborty, V. Nair, J. Konkel, S. Greene, M. Couper, C. Carlier, C. Wiese, R. Little, C. Pomerleau and O. Pomerleau, "The role of engagement in a tailored web-based smoking cessation

- program: randomized controlled trial," *J. Med. Internet Res.*, vol. 10, pp. e36, 5, 2008.
- [188] M. P. Couper, G.L. Alexander, N. Maddy, N. Zhang, M.A. Nowak, J.B. McClure, J.J. Calvi, S.J. Rolnick, M.A. Stopponi, R.J. Little, and C. Cole Johnson, "Engagement and retention: measuring breadth and depth of participant use of an online intervention," *J. Med. Internet Res.*, vol. 12, pp. e52, 4, 2010.
- [189] S. S. Anand and B. Mobasher, "Intelligent techniques for web personalization," in *Proceedings of the 2003 International Conference on Intelligent Techniques for Web Personalization*, 2003, pp. 1-36.
- [190] B. Mobasher, "Data mining for web personalization," in *The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds. Berlin: Springer, 2007, pp. 90-135.
- [191] S. Gallacher, E. Papadopoulou, Y. Abu-Shaaban, N.K. Taylor and M.H. Williams, "Dynamic context-aware personalisation in a pervasive environment," *Pervasive and Mobile Computing*, vol. 10, pp. 120-137, 2014.
- [192] S.S. Anand, P. Kearney and M. Shapcott, "Generating semantically enriched user profiles for web personalization," *ACM Transactions on Internet Technology (TOIT)*, vol. 7, pp. 22, 4, 2007.
- [193] T. Kuflik, J. Kay and B. Kummerfeld, "Challenges and Solutions of Ubiquitous User Modeling," in *Ubiquitous Display Environments*, A. Krüger and T. Kuflik, Eds. Berlin: Springer, 2012, pp. 7-30.
- [194] F. Abel, E. Herder, G. Houben, N. Henze and D. Krause, "Cross-system user modeling and personalization on the social web," *User Model. User Adapt. Interact.*, vol. 23, pp. 169-209, 2, 2013.
- [195] C. Emmanouilidis, R. Koutsiamanis and A. Tasidou, "Mobile guides: Taxonomy of architectures, context awareness, technologies and applications," *Journal of Network and Computer Applications*, vol. 36, pp. 103-125, 1, 2013.
- [196] A. L. Montgomery and M. D. Smith, "Prospects for Personalization on the Internet," *Journal of Interactive Marketing*, vol. 23, pp. 130-137, 5, 2009.
- [197] X. Zhou, Y. Xu, Y. Li, A. Josang and C. Cox, "The state-of-the-art in personalized recommender systems for social networking," *Artif. Intell. Rev.*, vol. 37, pp. 119-132, 2, 2012.

- [198] A. S. Das, M. Datar, A. Garg and S. Rajaram, "Google news personalization: Scalable online collaborative filtering," in *Proceedings of the 16th International Conference on World Wide Web*, 2007, pp. 271-280.
- [199] A. Sunikka and J. Bragge, "Applying text-mining to personalization and customization research literature—Who, what and where?" *Expert Syst. Appl.*, vol. 39, pp. 10049-10058, 11, 2012.
- [200] S. Gallacher, E. Papadopoulou, N.K. Taylor and M.H. Williams, "Learning user preferences for adaptive pervasive environments: an incremental and temporal approach," *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, vol. 8, pp. 5, 1, 2013.
- [201] E. Toch, Y. Wang and L. F. Cranor, "Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems," *User Model. User Adapt. Interact.*, vol. 22, pp. 203-220, 1-2, 2012.
- [202] D. Gavalas, C. Konstantopoulos, K. Mastakas and G. Pantziou, "Mobile recommender systems in tourism," *Journal of Network and Computer Applications*, vol. 39, pp. 319-333, 2014.
- [203] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Advances in Artificial Intelligence*, vol. 2009, 2009.
- [204] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowled. Data Eng.*, vol. 17, pp. 734-749, 6, 2005.
- [205] P. Lops, M. De Gemmis and G. Semeraro, "Content-based recommender systems: State of the art and trends," in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira and P. Kantor, Eds. Boston, MA: Springer, 2011, pp. 73-105.
- [206] A. Tiroshi, T. Kuflik, J. Kay and B. Kummerfeld, "Recommender systems and the social web," in *Advances in User Modeling. UMAP 2011, Lecture Notes in Computer Science, Vol. 4321, Vol. 7138*, L. Ardissono and T. Kuflik, Eds. Berlin: Springer, 2012, pp. 60-70.
- [207] D. H. Park, H.K. Kim, I.Y. Choi and J.K. Kim, "A literature review and classification of recommender systems research," *Expert Syst. Appl.*, vol. 39, pp. 10059-10072, 11, 2012.
- [208] A. C. Valdez, M. Ziefle, K. Verbert, A. Felfernig and A. Holzinger, "Recommender systems for health informatics: State-of-the-art and future

- perspectives," in *Machine Learning for Health Informatics*, A. Holzinger, Ed. Cham: Springer, 2016, pp. 391-414.
- [209] E. Sezgin and S. Ozkan, "A systematic literature review on health recommender systems," in *E-Health and Bioengineering Conference (EHB), 2013*, 2013, pp. 1-4.
- [210] H. Roitman, Y. Messika, Y. Tsimmerman and Y. Maman, "Increasing patient safety using explanation-driven personalized content recommendation," in *Proceedings of the 1st ACM International Health Informatics Symposium*, 2010, pp. 430-434.
- [211] H. Schäfer, S. Hors-Fraile, R.P. Karumur, A. Calero Valdez, A. Said, H. Torkamaan, T. Ulmer and C. Trattner, "Towards health (aware) recommender systems," in *Proceedings of the 2017 International Conference on Digital Health*, 2017, pp. 157-161.
- [212] C. L. Sanchez Bocanegra, J. L. Sevillano Ramos, C. Rizo, A. Civit and L. Fernandez-Luque, "HealthRecSys: A semantic content-based recommender system to complement health videos," *BMC Med. Inform. Decis. Mak.*, vol. 17, pp. 63, 2017.
- [213] D. Elswiler, M. Harvey, B. Ludwig and A. Said, "Bringing the "healthy" into food recommenders," in *Proceedings of the 2nd International Workshop on Decision Making and Recommender Systems, DMRS 2015*, 2015, pp. 33-36.
- [214] T. N. Trang Tran, M. Atas, A. Felfernig and M. Stettinger, "An overview of recommender systems in the healthy food domain," *J. Intell. Inf. Syst.*, pp. 1-26, 2017.
- [215] R. G. Farrell, C.M. Danis, S. Ramakrishnan and W.A. Kellogg, "Intrapersonal retrospective recommendation: Lifestyle change recommendations using stable patterns of personal behavior," in *Proceedings of the First International Workshop on Recommendation Technologies for Lifestyle Change (LIFESTYLE 2012)*, Dublin, Ireland, 2012, pp. 24.
- [216] T. Rist, A. Seiderer, S. Hammer, M. Mayr and E. André, "CARE: Extending a digital picture frame with a recommender mode to enhance well-being of elderly people," in *Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare*, 2015, pp. 112-120.

- [217] J. Treviranus and S. Hockema, "The value of the unpopular: Counteracting the popularity echo-chamber on the web," in *Science and Technology for Humanity (TIC-STH), 2009 IEEE Toronto International Conference*, 2009, pp. 603-608.
- [218] H. Ashman, T. Brailsford, A.I. Cristea, Q.Z. Sheng, C. Stewart, E.G. Toms and V. Wade, "The ethical and social implications of personalization technologies for e-learning," *Information & Management*, vol. 51, pp. 819-832, 6, 2014.
- [219] J. Passe, C. Drake and L. Mayger, "Homophily, echo chambers, & selective exposure in social networks: What should civic educators do?" *The Journal of Social Studies Research*, 2017.
- [220] A. Krasodonski-Jones, "Talking to Ourselves: Political debate and online echo chambers," Demos, London, 2016. [Online]. Available: <https://www.demos.co.uk/project/talking-to-ourselves/>. [Accessed: December 28, 2017].
- [221] H. T. Williams, J.R. McMurray, T. Kurz and F.H. Lambert, "Network analysis reveals open forums and echo chambers in social media discussions of climate change," *Global Environ. Change*, vol. 32, pp. 126-138, 2015.
- [222] E. Colleoni, A. Rozza and A. Arvidsson, "Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data," *J. Commun.*, vol. 64, pp. 317-332, 2, 2014.
- [223] A. Hannak, P. Sapiezynski, A. Molavi Kakhki, B. Krishnamurthy, D. Lazer, A. Mislove and C. Wilson, "Measuring personalization of web search," in *Proceedings of the 22nd International Conference on World Wide Web*, 2013, pp. 527-538.
- [224] H. Holone, "The filter bubble and its effect on online personal health information," *Croat. Med. J.*, vol. 57, pp. 298-301, 3, 2016.
- [225] A. Kobsa, "Supporting user interfaces for all through user modeling," *Advances in Human Factors/Ergonomics*, vol. 20, pp. 155-157, 1995.
- [226] E. Frias-martinez, S. Y. Chen and Xiaohui Liu, "Survey of Data Mining Approaches to User Modeling for Adaptive Hypermedia," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 36, pp. 734-749, 6, 2006.
- [227] L. Razmerita, "An Ontology-Based Framework for Modeling User Behavior—A Case Study in Knowledge Management," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, pp. 772-783, 4, 2011.

- [228] P. Brusilovsky and E. Millán, "User Models for Adaptive Hypermedia and Adaptive Educational Systems," in *The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds. Berlin: Springer, 2007, pp. 3-53.
- [229] F. Carmagnola, F. Cena and C. Gena, "User model interoperability: a survey," *User Model. User Adapt. Interact.*, vol. 21, pp. 285-331, 3, 2011.
- [230] A. Kavcic, "The role of user models in adaptive hypermedia systems," in *Electrotechnical Conference, 2000. MELECON 2000*, 2000, pp. 119-122.
- [231] S. Berkovsky, T. Kuflik and F. Ricci, "Cross-representation mediation of user models," *User Model. User Adapt. Interact.*, vol. 19, pp. 35-63, 1, 2009.
- [232] A. Kobsa, "Generic user modeling systems," in *The Adaptive Web: Methods and Strategies of Web Personalization. Lecture Notes in Computer Science, Vol. 4321*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds. Berlin: Springer, 2007, pp. 136-154.
- [233] J. Ye, S. Dobson and S. McKeever, "Situation identification techniques in pervasive computing: A review," *Pervasive and Mobile Computing*, vol. 8, pp. 36-66, 1, 2012.
- [234] X. Ding, Y. Li and Y. Zhao, "A framework of user model based on semi-supervised techniques," in *E-Business Engineering, 2008. ICEBE'08. IEEE International Conference on*, 2008, pp. 396-401.
- [235] P. S. Panchal and U. D. Agravat, "Hybrid technique for user's web page access prediction based on markov model," in *Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on*, 2013, pp. 1-8.
- [236] I. Zukerman and D. W. Albrecht, "Predictive statistical models for user modeling," *User Model. User Adapt. Interact.*, vol. 11, pp. 5-18, 1, 2001.
- [237] D. J. Cook and S. K. Das, "Pervasive computing at scale: Transforming the state of the art," *Pervasive and Mobile Computing*, vol. 8, pp. 22-35, 2, 2012.
- [238] J. A. Gulla and V. Sugumaran, Eds., *Applied Semantic Web Technologies*. Boca Raton: Auerbach Publications, 2011.
- [239] World Wide Web Consortium (2013, Dec). *Web Ontology Language (OWL)*. [Online]. Available: <https://www.w3.org/OWL/>. [Accessed: June 16, 2017].

- [240] D. Heckmann, E. Schwarzkopf, J. Mori, D. Dengler and A. Kröner, "The user model and context ontology gumo revisited for future web 2.0 extensions," in *Contexts and Ontologies: Representation and Reasoning (2007)*, 2007, pp. 37-46.
- [241] N. Ye, Y. Zhao, X. Zhao, E. Liu and G. Clapworthy, "A linked data-driven solution to user modeling in a multi-application environment," in *Software Engineering and Service Science (ICSESS), 2011 IEEE 2nd International Conference on*, 2011, pp. 872-876.
- [242] E. Walsh, A. O'Connor and V. Wade, "Evaluation of a domain-aware approach to user model interoperability," in *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, 2012, pp. 197-206.
- [243] M. Viviani, N. Bennani and E. Egyed-Zsigmond, "A survey on user modeling in multi-application environments," in *Advances in Human-Oriented and Personalized Mechanisms, Technologies and Services (CENTRIC), 2010 Third International Conference on*, 2010, pp. 111-116.
- [244] J. Fink and A. Kobsa, "A review and analysis of commercial user modeling servers for personalization on the world wide web," *User Model. User Adapt. Interact.*, vol. 10, pp. 209-249, 2, 2000.
- [245] C. Bettini, O. Brdiczka, K. Henriksen, J. Indulska, D. Nicklas, A. Ranganathan and D. Riboni, "A survey of context modelling and reasoning techniques," *Pervasive and Mobile Computing*, vol. 6, pp. 161-180, 2, 2010.
- [246] J. Hong, E. Suh and S. Kim, "Context-aware systems: A literature review and classification," *Expert Syst. Appl.*, vol. 36, pp. 8509-8522, 4, 2009.
- [247] A. K. Dey, "Understanding and using context," *Personal and Ubiquitous Computing*, vol. 5, pp. 4-7, 1, 2001.
- [248] A. Zimmermann, A. Lorenz and R. Oppermann, "An operational definition of context," in *Modeling and Using Context. CONTEXT 2007. Lecture Notes in Computer Science, Vol 4635*, B. Kokinov, D. Richardson, T. Roth-Berghofer and L. Vieu, Eds. Berlin: Springer, 2007, pp. 558-571.
- [249] B. Schilit, N. Adams and R. Want, "Context-aware computing applications," in *Mobile Computing Systems and Applications, 1994. WMCSA 1994. First Workshop on*, 1994, pp. 85-90.
- [250] C. Bolchini, C. A. Curino, E. Quintarelli, F. A. Schreiber and L. Tanca, "A data-oriented survey of context models," *ACM Sigmod Record*, vol. 36, pp. 19-26, 4, 2007.

- [251] L. Buriano, "Exploiting social context information in context-aware mobile tourism guides," in *Proc.of Mobile Guide 2006*, 2006.
- [252] J. E. Dickinson, K. Ghali, T. Cherrett, C. Speed, N. Davies and S. Norgate, "Tourism and the smartphone app: Capabilities, emerging practice and scope in the travel domain," *Current Issues in Tourism*, vol. 17, pp. 84-101, 1, 2014.
- [253] M. Migliardi and M. Gaudina, "A context aware, mobile system providing memory support for ageing people," in *MIPRO, 2011 Proceedings of the 34th International Convention*, 2011, pp. 535-540.
- [254] S. Najar, O. Saidani, M. Kirsch-Pinheiro, C. Souveyet and S. Nurcan, "Semantic representation of context models: A framework for analyzing and understanding," in *Proceedings of the 1st Workshop on Context, Information and Ontologies*, 2009, pp. 6.
- [255] M. Baldauf, S. Dustdar and F. Rosenberg, "A survey on context-aware systems," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 2, pp. 263-277, 4, 2007.
- [256] J. Ye, L. Coyle, S. Dobson and P. Nixon, "Ontology-based models in pervasive computing systems," *Knowl. Eng. Rev.*, vol. 22, pp. 315-347, 4, 2007.
- [257] T. Strang, C. Linnhoff-Popien, "A Context Modeling Survey," in *Proceedings of the First International Workshop on Advanced Context Modelling, Reasoning and Management, in conjunction with UbiComp 2004*, 2004.
- [258] A. Gómez-Pérez, M. Fernández-López and O. Corcho, "Ontological engineering: what are ontologies and how can we build them?" in *Semantic Web Services: Theory, Tools and Applications*, J. Cardoso, Ed. Hershey, PA: IGI Global, 2007, pp. 44-70.
- [259] M. Uschold and M. Gruninger, "Ontologies: Principles, methods and applications," *Knowl. Eng. Rev.*, vol. 11, pp. 93-136, 2, 1996.
- [260] T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing?" *International Journal of Human-Computer Studies*, vol. 43, pp. 907-928, 5-6, 1995.
- [261] N. F. Noy and D. L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," Stanford Knowledge Systems Laboratory, Stanford, CA, 2001. [Online]. Available: <http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html> [Accessed: June 16, 2016].

- [262] M. Horridge, "A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools Edition Edition 1.3.," The University Of Manchester, Manchester, 2011 [online] Available: http://mowl-power.cs.man.ac.uk/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf. [Accessed: June 16, 2017].
- [263] T. Berners-Lee, J. Hendler and O. Lassila, "The semantic web," *Sci. Am.*, vol. 284, pp. 28-37, 5, 2001.
- [264] H. P. Alesso and C. F. Smith, *Developing Semantic Web Services*. Natick, Mass.: A K Peters, 2005.
- [265] P. Hitzler, M. Krötzsch and S. Rudolph, *Foundations of Semantic Web Technologies*. London: Chapman & Hall/CRC, 2009.
- [266] G. Antoniou, P. Groth, F. van Harmelen and R. Hoekstra, *A Semantic Web Primer*. Cambridge, Mass.: MIT Press, 2012.
- [267] N. Shadbolt, T. Berners-Lee and W. Hall, "The semantic web revisited," *IEEE Intell. Syst.*, vol. 21, pp. 96-101, 3, 2006.
- [268] B. M. Konopka, "Biomedical ontologies—A review," *Biocybernetics and Biomedical Engineering*, vol. 35, pp. 75-86, 2, 2015.
- [269] R. Hoehndorf, M. Dumontier and G. V. Gkoutos, "Evaluation of research in biomedical ontologies," *Brief Bioinform.*, vol. 14, pp. 696-712, 6, 2013.
- [270] M. Ivanović and Z. Budimac, "An overview of ontologies and data resources in medical domains," *Expert Syst. Appl.*, vol. 41, pp. 5158-5166, 11, 2014.
- [271] World Wide Web Consortium (2014). *Resource Description Framework (RDF)*. [Online]. Available: <https://www.w3.org/RDF/>. [Accessed: June 16, 2017].
- [272] World Wide Web Consortium (2012, Dec). *OWL 2 Web Ontology Language Document Overview (Second Edition)*. [Online]. Available: <https://www.w3.org/TR/owl2-overview/>. [Accessed: December 19, 2017].
- [273] World Wide Web Consortium (2012, Dec). *OWL 2 Web Ontology Language Primer (Second Edition)*. [Online]. Available: <https://www.w3.org/TR/owl2-primer/>. [Accessed: December 19, 2017].
- [274] World Wide Web Consortium (2012, Dec). *OWL 2 Web Ontology Language Profiles (Second Edition)*. [Online]. Available: https://www.w3.org/TR/owl2-profiles/#OWL_2_QL. [Accessed: December 19, 2017].

- [275] S. Harris and A. Seaborne (2013, Mar). *SPARQL 1.1 Query Language*.
[Online]. Available: <https://www.w3.org/TR/sparql11-query/>. [Accessed: December 27, 2017].
- [276] World Wide Web Consortium (2017, Sep). *Semantic Web Tools*. [Online].
Available: <https://www.w3.org/2001/sw/wiki/SemanticWebTools>. [Accessed: December 19, 2017].
- [277] F. Zeshan and R. Mohamad, "Medical Ontology in the Dynamic Healthcare Environment," *Procedia Computer Science*, vol. 10, pp. 340-348, 2012.
- [278] S. Singh and R. Karwayun, "A comparative study of inference engines," in *Information Technology: New Generations (ITNG), 2010 Seventh International Conference on*, 2010, pp. 53-57.
- [279] T. Segaran, C. Evans and J. Taylor, *Programming the Semantic Web*. O'Reilly Media, 2009.
- [280] B. Glimm and H. Stuckenschmidt, "15 years of semantic web: An incomplete survey," *KI-Künstliche Intelligenz*, vol. 30, pp. 117-130, 2, 2016.
- [281] A. C. Yu, "Methods in biomedical ontology," *J. Biomed. Inform.*, vol. 39, pp. 252-266, 6, 2006.
- [282] O. Bodenreider, "Biomedical ontologies in action: role in knowledge management, data integration and decision support," *Yearb. Med. Inform.*, pp. 67-79, 2008.
- [283] A. Ruttenberg, T. Clark, W. Bug, M. Samwald, O. Bodenreider, H. Chen, D. Doherty, K. Forsberg, Y. Gao, V. Kashyap, J. Kinoshita, J. Luciano, M.S. Marshall, C. Ogbuji, J. Rees, S. Stephens, G.T. Wong, E. Wu, D. Zaccagnini, T. Hongsermeier, E. Neumann, I. Herman and K.H. Cheung, "Advancing translational research with the Semantic Web," *BMC Bioinformatics*, vol. 8 Suppl. 3, 2007.
- [284] R. Hoehndorf, P. N. Schofield and G. V. Gkoutos, "The role of ontologies in biological and biomedical research: a functional perspective," *Brief Bioinform.*, vol. 16, pp. 1069-1080, 6, 2015.
- [285] U.S. National Library of Medicine (2016, Apr). *Unified Medical Language System (UMLS)*. [Online]. Available: <https://www.nlm.nih.gov/research/umls/>. [Accessed: December 27, 2017].
- [286] OBO Technical WG (2017). *The OBO Foundry*. [Online]. Available: <http://www.obofoundry.org/>. [Accessed: December 27, 2017].

- [287] T. Narock and H. Wimmer, "Linked data scientometrics in semantic e-Science," *Comput. Geosci.*, vol. 100, pp. 87-93, 2017.
- [288] L. Ramos, "Semantic Web for manufacturing, trends and open issues: Toward a state of the art," *Comput. Ind. Eng.*, vol. 90, pp. 444-460, 2015.
- [289] E. Negri, L. Fumagalli, M. Garetti and L. Tanca, "Requirements and languages for the semantic representation of manufacturing systems," *Comput. Ind.*, vol. 81, pp. 55-66, 2016.
- [290] P. Pauwels, S. Zhang and Y. Lee, "Semantic web technologies in AEC industry: A literature overview," *Autom. Constr.*, vol. 73, pp. 145-165, 2017.
- [291] J. G. Breslin, D. O'Sullivan, A. Passant and L. Vasiliu, "Semantic Web computing in industry," *Comput. Ind.*, vol. 61, pp. 729-741, 8, 2010.
- [292] J. Rayfield (April, 2012). *Sports Refresh: Dynamic Semantic Publishing*. [Online]. Available: http://www.bbc.co.uk/blogs/bbcinternet/2012/04/sports_dynamic_semantic.html. [Accessed: December 19, 2017].
- [293] J. Rayfield, P. Haase and B. Popov (2012). *BBC Dynamic Semantic Publishing [DSP] SemtechBizUK 2012*. [Online]. Available: <https://speakerdeck.com/jemrayfield/bbc-dynamic-semantic-publishing-sport%7Colympics-semtechbiz-uk>. [Accessed: December 19, 2017].
- [294] M. Hepp (n.d.). *GoodRelations. The Web Vocabulary for E-Commerce*. [Online]. Available: <http://www.heppnetz.de/projects/goodrelations/>. [Accessed: December 19, 2017].
- [295] R. MacManus (July, 2010). *How Best Buy Is Using The Semantic Web*. [Online]. Available: <http://www.nytimes.com/external/readwriteweb/2010/07/01/01readwriteweb-how-best-buy-is-using-the-semantic-web-23031.html>. [Accessed: December 19, 2017].
- [296] S. Schulz and L. Jansen, "Formal ontologies in biomedical knowledge representation," *Yearb. Med. Inform.*, pp. 132-146, 2013.
- [297] S. Schulz, H. Stenzhorn, M. Boeker and B. Smith, "Strengths and limitations of formal ontologies in the biomedical domain," *Rev. Electron. Comun Inf. Inov Saude*, vol. 3, pp. 31-45, 1, 2009.
- [298] R. Stevens, M.E. Aranguren, K. Wolstencroft, U. Sattler, N. Drummond, M. Horridge and A. Rector, "Using OWL to model biological knowledge," *International Journal of Human-Computer Studies*, vol. 65, pp. 583-594, 7, 2007.

- [299] R. Hoehndorf, F. Loebe, J. Kelso and H. Herre, "Representing default knowledge in biomedical ontologies: Application to the integration of anatomy and phenotype ontologies," *BMC Bioinformatics*, vol. 8, pp. 377, 2007.
- [300] H. S. Pinto and J. P. Martins, "Ontologies: how can they be built?" *Knowledge and Information Systems*, vol. 6, pp. 441-464, 4, 2004.
- [301] O. Corcho, M. Fernández-López and A. Gómez-Pérez, "Methodologies, tools and languages for building ontologies. Where is their meeting point?" *Data Knowl. Eng.*, vol. 46, pp. 41-64, 1, 2003.
- [302] C. Soguero-Ruiz, L. Lechuga-Suárez, I. Mora-Jiménez, J. Ramos-López, Ó Barquero-Pérez, A. García-Alberola and J. L. Rojo-Álvarez, "Ontology for Heart Rate Turbulence Domain From The Conceptual Model of SNOMED-CT," *IEEE Transactions on Biomedical Engineering*, vol. 60, pp. 1825-1833, 7, 2013.
- [303] D. Isern, D. Sánchez and A. Moreno, "Ontology-driven execution of clinical guidelines," *Comput. Methods Programs Biomed.*, vol. 107, pp. 122-139, 8, 2012.
- [304] T. J. Bright, E. Yoko Furuya, G.J. Kuperman, J.J. Cimino and S. Bakken, "Development and evaluation of an ontology for guiding appropriate antibiotic prescribing," *J. Biomed. Inform.*, vol. 45, pp. 120-128, 2, 2012.
- [305] N. F. Noy, J. Mortensen, M.A. Musen and P.R. Alexander, "Mechanical turk as an ontology engineer?: Using microtasks as a component of an ontology-engineering workflow," in *Proceedings of the 5th Annual ACM Web Science Conference*, 2013, pp. 262-271.
- [306] Y. Ding and S. Foo, "Ontology research and development. Part 1-a review of ontology generation," *J. Inf. Sci.*, vol. 28, pp. 123-136, 2, 2002.
- [307] L. Zhou, "Ontology learning: state of the art and open issues," *Information Technology and Management*, vol. 8, pp. 241-252, 3, 2007.
- [308] T. Tudorache, N. F. Noy, S. Tu and M. A. Musen, "Supporting collaborative ontology development in protégé," in *The Semantic Web-ISWC 2008*, 2008, pp. 17-32.
- [309] M. A. Musen, N.F. Noy, N.H. Shah, P.L. Whetzel, C.G. Chute, M.A. Story, B. Smith and NCBO team, "The National Center for Biomedical Ontology," *J. Am. Med. Inform. Assoc.*, vol. 19, pp. 190-195, 2, 2012.
- [310] E. Simperl and M. Luczak-Rösch, "Collaborative ontology engineering: a survey," *Knowl. Eng. Rev.*, vol. 29, pp. 101-131, 1, 2014.

- [311] E. Simperl, "Reusing ontologies on the Semantic Web: A feasibility study," *Data Knowl. Eng.*, vol. 68, pp. 905-925, 10, 2009.
- [312] M. Strohmaier, S. Walk, J. Pöschko, D. Lamprecht, T. Tudorache, C. Nyulas, M. A. Musen and N. F. Noy, "How ontologies are made: Studying the hidden social dynamics behind collaborative ontology engineering projects," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 20, pp. 18-34, 5, 2013.
- [313] N. Kökciyan, R. Türkay, S. Üsküdarlı, P. Yolum, B. Bakir and B. Acar, "Semantic Description of Liver CT Images: An Ontological Approach," *IEEE J. Biomed. Health Inform.*, vol. 18, pp. 1363-1369, 4, 2014.
- [314] SNOMED International (2017). *What is SNOMED CT?*. [Online]. Available: <http://www.snomed.org/snomed-ct/what-is-snomed-ct>. [Accessed: June 16, 2017].
- [315] A. Ryan, "Towards semantic interoperability in healthcare: Ontology mapping from SNOMED-CT to HL7 version 3," in *Proceedings of the Second Australasian Workshop on Advances in Ontologies-Volume 72*, 2006, pp. 69-74.
- [316] C. Sáez, A. Bresó, J. Vicente, M. Robles and J.M. García-Gómez, "An HL7-CDA wrapper for facilitating semantic interoperability to rule-based Clinical Decision Support Systems," *Comput. Methods Programs Biomed.*, vol. 109, pp. 239-249, 3, 2013.
- [317] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosz and M. Dean (2004, May). *SWRL: A Semantic Web Rule Language Combining OWL and RuleML*. [Online]. Available: <https://www.w3.org/Submission/SWRL/>. [Accessed: June 16, 2017].
- [318] N. Laserra, A. Alesanco, S. Guillén and J. Garcia, "A three stage ontology-driven solution to provide personalized care to chronic patients at home," *J. Biomed. Inform.*, vol. 46, pp. 516-529, 3, 2013.
- [319] R. F. Navarro, M. D. Rodriguez and J. Favela, "Intervention Tailoring in Augmented Cognition Systems for Elders With Dementia," *IEEE J. Biomed. Health Inform.*, vol. 18, pp. 361-367, 1, 2014.
- [320] P. R. Gradie, M. Litster, R. Thomas, J. Vyas and M.R. Schiller, "SciReader enables reading of medical content with instantaneous definitions," *BMC Med. Inform. Decis. Mak.*, vol. 11, pp. 4, 2011.

- [321] J. Demelo, P. Parsons and K. Sedig, "Ontology-Driven Search and Triage: Design of a Web-Based Visual Interface for MEDLINE," *JMIR Med. Inform.*, vol. 5, pp. e4, 1, 2017.
- [322] A. Al-Busaidi, A. Gray and N. Fiddian, "Personalizing web information for patients: linking patient medical data with the web via a patient personal knowledge base," *Health. Informatics J.*, vol. 12, pp. 27-39, 1, 2006.
- [323] E. Rohrer, R. Motz and A. Díaz, "Ontology-based process for recommending health websites," in *Software Services for e-World. I3E 2010. IFIP Advances in Information and Communication Technology*, vol. 341., W. Cellary and E. Estevez, Eds. Berlin: Springer, 2010, pp. 205-214.
- [324] J. Niederdeppe, R.C. Hornik, B.J. Kelly, D.L. Frosch, A. Romantan, R.S. Stevens, F.K. Barg, J.L. Weiner and J.S. Schwartz, "Examining the dimensions of cancer-related information seeking and scanning behavior," *Health Commun.*, vol. 22, pp. 153-167, 2, 2007.
- [325] C. D. Norman and H. A. Skinner, "eHealth literacy: essential skills for consumer health in a networked world," *J. Med. Internet Res.*, vol. 8, pp. e9, 2, 2006.
- [326] C. Lagoe and D. Atkin, "Health anxiety in the digital age: An exploration of psychological determinants of online health information seeking," *Comput. Hum. Behav.*, vol. 52, pp. 484-491, 11, 2015.
- [327] A. M. Stone, A.M. Scott, S.C. Martin and D.E. Brashers, "Using information to manage uncertainty during organ transplantation," *Departures in Critical Qualitative Research*, vol. 2, pp. 42-60, 1, 2013.
- [328] D. E. Brashers, J.L. Neidig, S.M. Haas, L.K. Dobbs, L.W. Cardillo and J.A. Russell, "Communication in the management of uncertainty: The case of persons living with HIV or AIDS," *Communications Monographs*, vol. 67, pp. 63-84, 1, 2000.
- [329] P. K. Han, W. M. Klein and N. K. Arora, "Varieties of uncertainty in health care: a conceptual taxonomy," *Med. Decis. Making*, vol. 31, pp. 828-838, 6, 2011.
- [330] D. E. Brashers, "Communication and uncertainty management," *J. Commun.*, vol. 51, pp. 477-497, 3, 2001.
- [331] S. C. Martin, A.M. Stone, A.M. Scott and D.E. Brashers, "Medical, personal, and social forms of uncertainty across the transplantation trajectory," *Qual. Health Res.*, vol. 20, pp. 182-196, 2, 2010.

- [332] F. Opreescu, S. Campo, J. Lowe, J. Andsager and J.A. Morcuende, "Online information exchanges for parents of children with a rare health condition: key findings from an online support community," *J. Med. Internet Res.*, vol. 15, pp. e16, 1, 2013.
- [333] A. M. Stone and C. L. Jones, "Sources of uncertainty: Experiences of Alzheimer's disease," *Issues Ment. Health Nurs.*, vol. 30, pp. 677-686, 11, 2009.
- [334] N. N. Vevea and A. N. Miller, "Patient Narratives: Exploring the Fit of Uncertainty-Management Models of Health Care," *Rev. Commun.*, vol. 10, pp. 276-289, 4, 2010.
- [335] D. E. DeLorme and J. Huh, "Seniors' uncertainty management of direct-to-consumer prescription drug advertising usefulness," *Health Commun.*, vol. 24, pp. 494-503, 6, 2009.
- [336] N. M. Thompson, J. L. Bevan and L. Sparks, "Healthcare reform information-seeking: Relationships with uncertainty, uncertainty discrepancy, and health self-efficacy," *Journal of Communication in Healthcare*, vol. 5, pp. 56-66, 1, 2012.
- [337] E. Donovan-Kicken and J. J. Bute, "Uncertainty of social network members in the case of communication-debilitating illness or injury," *Qual. Health Res.*, vol. 18, pp. 5-18, 1, 2008.
- [338] S. A. Rains, "Health information seeking and the world wide web: An uncertainty management perspective," *J. Health Commun.*, vol. 19, pp. 1296-1307, 11, 2014.
- [339] S. A. Rains and R. Tukachinsky, "An examination of the relationships among uncertainty, appraisal, and information-seeking behavior proposed in uncertainty management theory," *Health Commun.*, vol. 30, pp. 339-349, 4, 2015.
- [340] S. A. Rains and R. Tukachinsky, "Information Seeking in Uncertainty Management Theory: Exposure to Information About Medical Uncertainty and Information-Processing Orientation as Predictors of Uncertainty Management Success," *J. Health Commun.*, vol. 20, pp. 1275-1286, 11, 2015.
- [341] J. B. Barbour, L.S. Rintamaki, J.A. Ramsey and D.E. Brashers, "Avoiding health information," *J. Health Commun.*, vol. 17, pp. 212-229, 2, 2012.
- [342] D. E. Brashers, J. L. Neidig and D. J. Goldsmith, "Social support and the management of uncertainty for people living with HIV or AIDS," *Health Commun.*, vol. 16, pp. 305-331, 3, 2004.

- [343] G. L. Kreps and L. Neuhauser, "New directions in eHealth communication: Opportunities and challenges," *Patient Educ. Couns.*, vol. 78, pp. 329-336, 3, 2010.
- [344] R. P. Joseph, N.H. Durant, T.J. Benitez and D.W. Pekmezi, "Internet-Based Physical Activity Interventions," *Am. J. Lifestyle Med.*, vol. 8, pp. 42-67, 1, 2014.
- [345] L. Neuhauser and G. L. Kreps, "eHealth communication and behavior change: promise and performance," *Social Semiotics*, vol. 20, pp. 9-27, 1, 2010.
- [346] C. Robinson and J. Graham, "Perceived Internet health literacy of HIV-positive people through the provision of a computer and Internet health education intervention," *Health Info. Libr. J.*, vol. 27, pp. 295-303, 4, 2010.
- [347] B. Xie, "Improving older adults' e-health literacy through computer training using NIH online resources," *Libr. Inf. Sci. Res.*, vol. 34, pp. 63-71, 1, 2012.
- [348] S. Mitsutake, A. Shibata, K. Ishii and K. Oka, "Association of eHealth literacy with colorectal cancer knowledge and screening practice among internet users in Japan," *J. Med. Internet Res.*, vol. 14, pp. e153, 6, 2012.
- [349] H. Park and E. Lee, "Self-reported eHealth literacy among undergraduate nursing students in South Korea: A pilot study," *Nurse Educ. Today*, vol. 35, pp. 408-413, 2, 2015.
- [350] A. Tubaishat and L. Habiballah, "eHealth literacy among undergraduate nursing students," *Nurse Educ. Today*, vol. 42, pp. 47-52, 7, 2016.
- [351] E. Neter and E. Brainin, "eHealth literacy: extending the digital divide to the realm of health information," *J. Med. Internet Res.*, vol. 14, pp. e19, 1, 2012.
- [352] R. K. Britt and K. N. Hatten, "Need for Cognition and Electronic Health Literacy and Subsequent Information Seeking Behaviors Among University Undergraduate Students," *SAGE Open*, vol. 3, 4, 2013.
- [353] C. Knapp, V. Madden, M. Marcu, H. Wang, C. Curtis, P. Sloyer and E. Shenkman, "Information seeking behaviors of parents whose children have life-threatening illnesses," *Pediatric Blood & Cancer*, vol. 56, pp. 805-811, 5, 2011.
- [354] C. Knapp, V. Madden, H. Wang, P. Sloyer and E. Shenkman, "Internet use and eHealth literacy of low-income parents whose children have special health care needs," *J. Med. Internet Res.*, vol. 13, pp. e75, 3, 2011.
- [355] M. Stellefson, B. Hanik, J.D. Chaney and B. Tennant, "Analysis of ehealth search perspectives among female college students in the health professions using Q methodology," *J. Med. Internet Res.*, vol. 14, pp. e60, 2, 2012.

- [356] L. Ivanitskaya, I. O. Boyle and A. M. Casey, "Health information literacy and competencies of information age students: results from the interactive online Research Readiness Self-Assessment (RRSA)," *J. Med. Internet Res.*, vol. 8, pp. e6, 2, 2006.
- [357] A. J. A. M. van Deursen and J. A. G. M. van Dijk, "Using the Internet: Skill related problems in users' online behavior," *Interact. Comput.*, vol. 21, pp. 393-402, 12, 2009.
- [358] B. Tennant, M. Stellefson, V. Dodd, B. Chaney, D. Chaney, S. Paige and J. Alber, "eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults," *J. Med. Internet Res.*, vol. 17, pp. e70, 3, 2015.
- [359] B. D. Weiss, M.Z. Mays, W. Martz, K.M. Castro, D.A. DeWalt, M.P. Pignone, J. Mockbee and F.A. Hale, "Quick assessment of literacy in primary care: the newest vital sign," *Ann. Fam. Med.*, vol. 3, pp. 514-522, 6, 2005.
- [360] C. D. Norman and H. A. Skinner, "eHEALS: the eHealth literacy scale," *J. Med. Internet Res.*, vol. 8, pp. e27, 4, 2006.
- [361] Crown copyright (2016, Oct. 25). *What qualification levels mean*. [Online]. Available: <https://www.gov.uk/what-different-qualification-levels-mean/list-of-qualification-levels>. [Accessed: June 16, 2017].
- [362] G. Rowlands, N. Khazaezadeh, E. Oteng-Ntim, P. Seed, S. Barr and B.D. Weiss, "Development and validation of a measure of health literacy in the UK: the newest vital sign," *BMC Public Health*, vol. 13, pp. 1, 2013.
- [363] R. Capra, "HCI browser: A tool for studying web search behavior," *Proc. Am. Soc. Info. Sci. Tech.*, vol. 47, pp. 1-2, 1, 2010.
- [364] Business Services Organisation. (n.d.). *Prescribing by GP Practice*. [Online]. Available: <http://www.hscbusiness.hscni.net/services/2471.htm>. [Accessed: June 16, 2017].
- [365] World Health Organization (2017). *Diabetes country profiles 2016*. [Online]. Available: <http://www.who.int/diabetes/country-profiles/en/>. [Accessed: June 16, 2017].
- [366] World Health Organization (2017). *Obesity and overweight*. [Online]. Available: <http://www.who.int/mediacentre/factsheets/fs311/en/>. [Accessed: June 16, 2017].

- [367] Diabetes UK (2017). *EYES (RETINOPATHY)*. [Online]. Available: <https://www.diabetes.org.uk/Guide-to-diabetes/Complications/Retinopathy/>. [Accessed: June 16, 2017].
- [368] Public Health England, "Government 5 A Day logo licensing guidelines," Public Health England, London, 2016. [Online]. Available: <https://www.gov.uk/government/publications/government-5-a-day-logo>. [Accessed: June 16, 2017].
- [369] Department of Health (2015). *Antibiotics*. [Online]. Available: <http://www.nhs.uk/Conditions/Antibiotics-penicillins/Pages/Introduction.aspx>. [Accessed: June 16, 2017].
- [370] World Health Organization (2016, Nov). *Diabetes*. [Online]. Available: <http://www.who.int/mediacentre/factsheets/fs312/en/>. [Accessed: June 16, 2017].
- [371] R. van der Vaart, A.J. van Deursen, C.H. Drossaert, E. Taal, J.A. van Dijk and M. van de Laar, "Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations," *J. Med. Internet Res.*, vol. 13, pp. e86, 4, 2011.
- [372] J. Aponte and K. M. Nokes, "Electronic health literacy of older Hispanics with diabetes," *Health. Promot. Int.*, vol. 32, pp. 482-489, 3, 2017.
- [373] E. A. Chavarria, E.H. Chaney, M.L. Stellefson, J.D. Chaney, N. Chavarria and V.J. Dodd, "Types and Factors Associated With Online Health Information Seeking Among College Men in Latino Fraternities: A Qualitative Study," *Am. J. Mens. Health*, 2016. doi: <https://doi.org/10.1177/1557988315626510>.
- [374] P. Best, E. Gil-Rodriguez, R. Manktelow and B.J. Taylor, "Seeking Help From Everyone and No-One: Conceptualizing the Online Help-Seeking Process Among Adolescent Males," *Qual. Health Res.*, vol. 26, pp. 1067-1077, 8, 2016.
- [375] P. C. Pang, K. Verspoor, S. Chang and J. Pearce, "Conceptualising health information seeking behaviours and exploratory search: result of a qualitative study," *Health and Technology*, vol. 5, pp. 45-55, 1, 2015.
- [376] I. Rootman and D. Gordon-El-Bihbety, "A Vision for a Health Literate Canada. Report of the Expert Panel on Health Literacy," Canadian Public Health Association, Ottawa, 2008. [Online]. Available: https://www.cpha.ca/sites/default/files/uploads/resources/healthlit/report_e.pdf. [Accessed: January 10, 2018].

- [377] X. Sun, Y. Shi, Q. Zeng, Y. Wang, W. Du, N. Wei, R. Xie and C. Chang, "Determinants of health literacy and health behavior regarding infectious respiratory diseases: a pathway model," *BMC Public Health*, vol. 13, pp. 261, 2013.
- [378] W. H. Dutton and G. Blank with D. Groselj, "Cultures of the Internet: The Internet in Britain. Oxford Internet Survey 2013," Oxford Internet Institute, Oxford, 2013. [Online]. Available: <http://oxis.oii.ox.ac.uk/wp-content/uploads/sites/43/2014/11/OxIS-2013.pdf>. [Accessed: January 10, 2018].
- [379] Pew Research Center (2017, Jan). *Internet/Broadband Fact Sheet*. [Online]. Available: <http://www.pewinternet.org/fact-sheet/internet-broadband/>. [Accessed: January 10, 2018].
- [380] A. Perrin, "Social Media Usage: 2005-2015," Pew Research Center, Washington, DC, 2015. [Online]. Available: <http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/>. [Accessed: January 10, 2018].
- [381] Eurostat (2017, Dec). *Internet access and use statistics - households and individuals*. [Online]. Available: http://ec.europa.eu/eurostat/statistics-explained/index.php/Internet_access_and_use_statistics_-_households_and_individuals#Main_statistical_findings. [Accessed: January 10, 2018].
- [382] J. Berthelot, B. Le Goff and Y. Maugars, "The Hawthorne effect: Stronger than the placebo effect?" *Joint Bone Spine*, vol. 78, pp. 335-336, 4, 2011.
- [383] K.L. Leonard, "Is patient satisfaction sensitive to changes in the quality of care? An exploitation of the Hawthorne effect," *Journal of Health Economics*, vol. 27, pp. 444-459, 2008.
- [384] G. Seçkin, D. Yeatts, S. Hughes, C. Hudson and V. Bell, "Being an Informed Consumer of Health Information and Assessment of Electronic Health Literacy in a National Sample of Internet Users: Validity and Reliability of the e-HLS Instrument," *J. Med. Internet Res.*, vol. 18, pp. e161, 7, 2016.
- [385] R. van der Vaart and C. Drossaert, "Development of the Digital Health Literacy Instrument: Measuring a Broad Spectrum of Health 1.0 and Health 2.0 Skills," *J. Med. Internet Res.*, vol. 19, pp. e27, 1, 2017.

- [386] J. DeMarco and M. Nystrom, "The Importance of Health Literacy in Patient Education," *Journal of Consumer Health on the Internet*, vol. 14, pp. 294-301, 3, 2010.
- [387] International Diabetes Federation, "IDF Diabetes Atlas - Seventh Edition," International Diabetes Federation, Brussels, 2015. [Online]. Available: <http://www.diabetesatlas.org/>. [Accessed: June 16, 2017].
- [388] Diabetes UK (2017). *Diabetes UK. Know Diabetes. Fight Diabetes*. [Online]. Available: <https://www.diabetes.org.uk/>. [Accessed: June 16 2017].
- [389] American Diabetes Association (2017). *American Diabetes Association*. [Online]. Available: <http://www.diabetes.org/>. [Accessed: June 16 2017].
- [390] J. McDowell, F. Brown and D. Matthews, Eds., *Diabetes A Handbook for the Primary Healthcare Team*. Edinburgh: Churchill Livingstone, 2007.
- [391] F. Al Sayah and B. Williams, "An Integrated Model of Health Literacy Using Diabetes as an Exemplar," *Canadian Journal of Diabetes*, vol. 36, pp. 27-31, 2, 2012.
- [392] A. Coppola, L. Sasso, A. Bagnasco, A. Giustina and C. Gazzaruso, "The role of patient education in the prevention and management of type 2 diabetes: an overview," *Endocrine*, vol. 53, pp. 18-27, 1, 2016.
- [393] U. Sarkar, A.J. Karter, J.Y. Liu, H.H. Moffet, N.E. Adler and D. Schillinger, "Hypoglycemia is more common among type 2 diabetes patients with limited health literacy: the Diabetes Study of Northern California (DISTANCE)," *Journal of General Internal Medicine*, vol. 25, pp. 962-968, 9, 2010.
- [394] J. M. Mancuso, "Impact of health literacy and patient trust on glycemic control in an urban USA population," *Nurs. Health Sci.*, vol. 12, pp. 94-104, 1, 2010.
- [395] M. Clark, "Diabetes self-management education: A review of published studies," *Primary Care Diabetes*, vol. 2, pp. 113-120, 9, 2008.
- [396] C. Duncan, C. Birkmeyer, S. Coughlin, Q.E. Li, D. Sherr and S. Boren, "Assessing the value of diabetes education," *Diabetes Educ.*, vol. 35, pp. 752-60, 5, 2009.
- [397] D. Pacaud, H. Kelley, A.M. Downey and M. Chiasson, "Successful delivery of diabetes self-care education and follow-up through eHealth media," *Canadian Journal of Diabetes*, vol. 36, pp. 257-262, 5, 2012.
- [398] M. Trento, S. Gamba, L. Gentile, G. Grassi, V. Miselli, G. Morone, P. Passera, L. Tonutti, M. Tomalino, P. Bondonio, F. Cavallo, M. Porta and ROMEO

- Investigators, "Rethink Organization to iMprove Education and Outcomes (ROMEO): a multicenter randomized trial of lifestyle intervention by group care to manage type 2 diabetes," *Diabetes Care*, vol. 33, pp. 745-747, 4, 2010.
- [399] C. A. Chrvala, D. Sherr and R. D. Lipman, "Diabetes self-management education for adults with type 2 diabetes mellitus: A systematic review of the effect on glycemic control," *Patient Educ. Couns.*, vol. 99, pp. 926–943, 6, 2016.
- [400] B. Jankowska-Polańska, A.M. Fal, I. Uchmanowicz, M. Seń, J. Polański and D. Kurpas, "Influence of organized diabetic education on self-control and quality of life of patients with type 2 diabetes," *International Journal of Diabetes in Developing Countries*, vol. 35 Suppl. 2, pp. 79-87, 2015.
- [401] J. Lawton and D. Rankin, "How do structured education programmes work? An ethnographic investigation of the dose adjustment for normal eating (DAFNE) programme for type 1 diabetes patients in the UK," *Soc. Sci. Med.*, vol. 71, pp. 486-493, 8, 2010.
- [402] A. Steinsbekk, L. Rygg, M. Lisulo, M.B. Rise and A. Fretheim, "Group based diabetes self-management education compared to routine treatment for people with type 2 diabetes mellitus. A systematic review with meta-analysis," *BMC Health Serv Res.*, vol. 12, pp. 1, 2012.
- [403] S. Kim, F. Love, D.A. Quistberg and J.A. Shea, "Association of health literacy with self-management behavior in patients with diabetes," *Diabetes Care*, vol. 27, pp. 2980-2982, 12, 2004.
- [404] F. Hill-Briggs and A. S. Smith, "Evaluation of diabetes and cardiovascular disease print patient education materials for use with low-health literate populations," *Diabetes Care*, vol. 31, pp. 667-671, 4, 2008.
- [405] H. Y. Catherine, J.A. Parsons, S. Hall, D. Newton, A. Jovicic, D. Lottridge, B.R. Shah and S.E. Straus, "User-centered design of a web-based self-management site for individuals with type 2 diabetes-providing a sense of control and community," *BMC Med. Inform. Decis. Mak.*, vol. 14, pp. 60, 2014.
- [406] Y. H. Tang, S. Pang, M.F. Chan, G.S. Yeung and V.T. Yeung, "Health literacy, complication awareness, and diabetic control in patients with type 2 diabetes mellitus," *J. Adv. Nurs.*, vol. 62, pp. 74-83, 1, 2008.
- [407] A. Valls, K. Gibert, D. Sánchez and M. Batet, "Using ontologies for structuring organizational knowledge in Home Care assistance," *Int. J. Med. Inform.*, vol. 79, pp. 370-387, 5, 2010.

- [408] E. Sirin, B. Parsia, B.C. Grau, A. Kalyanpur and Y. Katz, "Pellet: A practical OWL-DL reasoner," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 5, pp. 51-53, 2, 2007.
- [409] C. L. Jackson, S. Bolen, F.L. Brancati, M.L. Batts-Turner and T.L. Gary, "A systematic review of interactive computer-assisted technology in diabetes care," *Journal of General Internal Medicine*, vol. 21, pp. 105-110, 2, 2006.
- [410] K. Pal, S.V. Eastwood, S. Michie, A.J. Farmer, M.L. Barnard, R. Peacock, B. Wood, J.D. Inniss and E. Murray, "Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus," *Cochrane Database of Systematic Reviews*, Issue 3, 2013. doi: 10.1002/14651858.CD008776.pub2.
- [411] M. Song, M. Choe, K.S. Kim, M.S. Yi, I. Lee, J. Kim, M. Lee, Y.M. Cho and Y.S. Shim, "An evaluation of Web-based education as an alternative to group lectures for diabetes self-management," *Nurs. Health Sci.*, vol. 11, pp. 277-284, 3, 2009.
- [412] S. A. Boren, T.L. Gunlock, M.M. Peebles and S. Krishna, "Computerized learning technologies for diabetes: a systematic review," *J. Diabetes Sci. Technol.*, vol. 2, pp. 139-146, 1, 2008.
- [413] K. Lorig, P.L. Ritter, D.D. Laurent, K. Plant, M. Green, V.B. Jernigan and S. Case, "Online diabetes self-management program: a randomized study," *Diabetes Care*, vol. 33, pp. 1275-1281, 6, 2010.
- [414] R. E. Glasgow, D. Kurz, D. King, J.M. Dickman, A.J. Faber, E. Halterman, T. Woolley, D.J. Toobert, L.A. Strycker and P.A. Estabrooks, "Twelve-month outcomes of an Internet-based diabetes self-management support program," *Patient Educ. Couns.*, vol. 87, pp. 81-92, 1, 2012.
- [415] P. C. Tang, J.M. Overhage, A.S. Chan, N.L. Brown, B. Aghighi, M.P. Entwistle, S.L. Hui, S.M. Hyde, L.H. Klieman, C.J. Mitchell, A.J. Perkins, L.S. Qureshi, T.A. Waltmyer, L.J. Winters and C.Y. Young, "Online disease management of diabetes: engaging and motivating patients online with enhanced resources-diabetes (EMPOWER-D), a randomized controlled trial," *J. Am. Med. Inform. Assoc.*, vol. 20, pp. 526-534, 3, 2013.
- [416] M.A. Orgun and T. Meyer, "Introduction to the special issue on advances in ontologies," *Expert Syst.*, vol. 25, pp. 175-178, 3, 2008.

- [417] A. Arwan, M. Sidiq, B. Priyambadha, H. Kristianto and R. Sarno, "Ontology and semantic matching for diabetic food recommendations," in *2013 International Conference on Information Technology and Electrical Engineering (ICITEE)*, 2013, pp. 170-175.
- [418] R. Chen, Y. Huang, C. Bau and S. Chen, "A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection," *Expert Syst. Appl.*, vol. 39, pp. 3995-4006, 3, 2012.
- [419] N. Chammas, R. Juric, N. Koay, V. Gurupur and S. C. Suh, "Towards a software tool for raising awareness of diabetic foot in diabetic patients," in *2013 46th Hawaii International Conference on System Sciences*, 2013, pp. 2646-2655.
- [420] A. Rahimi, S. Liaw, J. Taggart, P. Ray and H. Yu, "Validating an ontology-based algorithm to identify patients with Type 2 Diabetes Mellitus in Electronic Health Records," *Int. J. Med. Inform.*, vol. 83, pp. 768-778, 10, 2014.
- [421] R. F. Alharbi, J. Berri and S. El-Masri, "Ontology based clinical decision support system for diabetes diagnostic," in *2015 Science and Information Conference (SAI)*, 2015, pp. 597-602.
- [422] R. Johnson, J. Hoeller, K. Donald, C. Sampaleanu, R. Harrop, T. Risberg, A. Arendsen, D. Davison, D. Kopylenko, M. Pollack, T. Templier, E. Vervaet, P. Tung, B. Hale, A. Colyer, J. Lewis, C. Leau, M. Fisher, S. Brannen, R. Laddad, A. Poutsma, C. Beams, T. Abedrabbo, A. Clement, D. Syer, O. Gierke, R. Stoyanchev, P. Webb, R. Winch, B. Clozel, S. Nicoll and S. Deleuze (2016). *Web MVC framework*. [Online]. Available: <https://docs.spring.io/spring/docs/current/spring-framework-reference/html/mvc.html>. [Accessed: June 16, 2017].
- [423] Oracle. (n.d.). *JavaServer Pages Technology*. [Online]. Available: <http://www.oracle.com/technetwork/java/javaee/jsp/index.html>. [Accessed: June 16, 2017].
- [424] L. Al-Jadir, C. Parent and S. Spaccapietra, "Reasoning with large ontologies stored in relational databases: The OntoMinD approach," *Data Knowl. Eng.*, vol. 69, pp. 1158-1180, 11, 2010.
- [425] International Diabetes Federation (2017). *Promoting diabetes care, prevention and a cure worldwide*. [Online]. Available: <http://www.idf.org>. [Accessed: June 28, 2017].

- [426] Stanford Center for Biomedical Informatics Research (2016). *Protégé*. [Online]. Available: <http://protege.stanford.edu/>. [Accessed: June 16, 2017].
- [427] Department of Health and Social Care, "Start Active, Stay Active," Crown copyright, 2011. [Online]. Available: <https://www.gov.uk/government/publications/start-active-stay-active-a-report-on-physical-activity-from-the-four-home-countries-chief-medical-officers>. [Accessed: January 14, 2018].
- [428] M. O'Connor (2017, May). *SWRLLanguageFAQ*. [Online]. Available: <https://github.com/protegeproject/swrlapi/wiki/SWRLLanguageFAQ>. [Accessed: June 16, 2017].
- [429] National Institute of Adult Continuing Education, "Readability. How to produce clear written materials for a range of readers," National Institute of Adult Continuing Education, London, 2009. [Online]. Available: <http://shop.niace.org.uk/readability.html>. [Accessed: Jan. 12, 2017].
- [430] Department for Education and Skills, Qualifications and Curriculum Authority and The Basic Skills Agency, "National standards for adult literacy, numeracy and ICT"" Qualifications and Curriculum Authority, London, 2005. [Online]. Available: https://set.foundation.co.uk/media/131246/2005_national_standards_for_adult_literacy_numeracy_ict.pdf. [Accessed: June 16, 2017].
- [431] Quality Assurance Agency for Higher Education, Office of Qualifications and Examinations Regulation, Scottish Credit and Qualifications Framework Partnership, Council for the Curriculum Examinations and Assessment Accreditation, Credit and Qualifications Framework for Wales, Quality and Qualifications Ireland, "Qualifications can Cross Boundaries - A Guide to Comparing Qualifications in the UK and Ireland. Sixth Edition," 2017. [Online]. Available: http://www.qaa.ac.uk/publications/information-and-guidance/publication?PubID=165#.WHbs_H3eJSV. [Accessed: June 16, 2017].
- [432] Oracle (n.d.). *Java 8 Central*. [Online]. Available: <http://www.oracle.com/technetwork/java/javase/overview/java8-2100321.html>. [Accessed: June 16, 2017].
- [433] University of Manchester. (n.d.). *The OWL API*. [Online]. Available: <http://owlapi.sourceforge.net/>. [Accessed: June 16, 2017].

- [434] Oracle Corporation. (2017). *MySQL*. [Online]. Available: <https://www.mysql.com/>. [Accessed: June 16, 2017].
- [435] R. Johnson, J. Hoeller, K. Donald, C. Sampaleanu, R. Harrop, T. Risberg, A. Arendsen, D. Davison, D. Kopylenko, M. Pollack, T. Templier, E. Vervaet, P. Tung, B. Hale, A. Colyer, J. Lewis, C. Leau, M. Fisher, S. Brannen, R. Laddad, A. Poutsma, C. Beams, T. Abedrabbo, A. Clement, D. Syer, O. Gierke, R. Stoyanchev, P. Webb, R. Winch, B. Clozel, S. Nicoll and S. Deleuze (2016). *Data access with JDBC*. [Online]. Available: <https://docs.spring.io/spring/docs/current/spring-framework-reference/html/jdbc.html>. [Accessed: June 16, 2017].
- [436] The Apache Software Foundation. (2017). *Apache Tomcat*. [Online]. Available: <http://tomcat.apache.org/>. [Accessed: June 16, 2017].
- [437] Diabetes UK, "Your Guide To Type 2 Diabetes," Diabetes UK, London, 2015. [Online]. Available: <http://cepn.barnetccg.nhs.uk/Downloads/Diabetes%20UK%20Type%20%20infor%20mation%20booklet.pdf>. [Accessed: January 14, 2018].
- [438] Diabetes.co.uk (2018). *Diabetes.co.uk the global diabetes community*. [Online]. Available: <https://www.diabetes.co.uk/>. [Accessed: January 10, 2018].
- [439] O. Bodenreider and R. Stevens, "Bio-ontologies: current trends and future directions," *Brief Bioinform.*, vol. 7, pp. 256-274, 3, 2006.
- [440] J. Brank, M. Grobelnik and D. Mladenic, "A survey of ontology evaluation techniques," in *Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005)*, 2005, pp. 166-170.
- [441] A. Gómez-Pérez, "Evaluation of taxonomic knowledge in ontologies and knowledge bases," in *Proceedings of the Banff Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW'99)*, 1999, pp. 6.1.1 - 6.1.18.
- [442] L. Obrst, C. Werner, M. Inderjeet, S. Ray and B. Smith, "The Evaluation of Ontologies: Toward Improved Semantic Interoperability," in *Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences*, C. Baker and K. H. Cheung, Eds. New York: Springer, 2006, pp. 139-158.
- [443] P. Delir Haghighi, F. Burstein, A. Zaslavsky and P. Arbon, "Development and evaluation of ontology for intelligent decision support in medical emergency management for mass gatherings," *Decis. Support Syst.*, vol. 54, pp. 1192-1204, 1, 2013.

- [444] C. L. Gordon and C. Weng, "Combining expert knowledge and knowledge automatically acquired from electronic data sources for continued ontology evaluation and improvement," *J. Biomed. Inform.*, vol. 57, pp. 42-52, 10, 2015.
- [445] Arizona State University (n.d.). *Compendium of Physical Activities*. [Online]. Available: <https://sites.google.com/site/compendiumofphysicalactivities/home>. [Accessed: June 16, 2017].
- [446] K. Dramé, G. Diallo, F. Delva, J.F. Dartigues, E. Mouillet, R. Salamon and F. Mougin, "Reuse of termino-ontological resources and text corpora for building a multilingual domain ontology: An application to Alzheimer's disease," *J. Biomed. Inform.*, vol. 48, pp. 171-182, 4, 2014.
- [447] B. Hu, S. Dasmahapatra, D. Dupplaw, P. Lewis and N. Shadbolt, "Reflections on a medical ontology," *International Journal of Human-Computer Studies*, vol. 65, pp. 569-582, 7, 2007.
- [448] R.B. Trelease, "Anatomical reasoning in the informatics age: Principles, ontologies, and agendas," *The Anatomical Record*, vol. 289B, pp. 72-84, 2, 2006.
- [449] J. Hong, T. V. Nguyen and N. S. Prose, "Compassionate care: Enhancing physician-patient communication and education in dermatology: Part II: Patient education," *J. Am. Acad. Dermatol.*, vol. 68, pp. 364.e1-364.e10, 3, 2013.
- [450] A. Colledge, J. Car, A. Donnelly and A. Majeed, "Health information for patients: time to look beyond patient information leaflets," *J. R. Soc. Med.*, vol. 101, pp. 447-453, 9, 2008.
- [451] W. Kroeze, A. Werkman and J. Brug, "A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors," *Ann. Behav. Med.*, vol. 31, pp. 205-223, 3, 2006.
- [452] L. Attack and R. Luke, "The impact of validated, online health education resources on patient and community members' satisfaction and health behaviour," *Health Educ. J.*, vol. 71, pp. 211-218, 2, 2012.
- [453] N. B. Giuse, T.Y. Koonce, A.B. Storrow, S.V. Kusnoor and F. Ye, "Using health literacy and learning style preferences to optimize the delivery of health information," *J. Health Commun.*, vol. 17 Suppl. 3, pp. 122-140, 2012.
- [454] H. L. O'Brien and E. G. Toms, "What is user engagement? A conceptual framework for defining user engagement with technology," *J. Am. Soc. Inf. Sci. Technol.*, vol. 59, pp. 938-955, 6, 2008.

- [455] H. L. O'Brien and E. G. Toms, "Examining the generalizability of the User Engagement Scale (UES) in exploratory search," *Information Processing & Management*, vol. 49, pp. 1092-1107, 5, 2013.
- [456] H. O'Brien and P. Cairns, "An empirical evaluation of the User Engagement Scale (UES) in online news environments," *Information Processing & Management*, vol. 51, pp. 413-427, 4, 2015.
- [457] M. B. Cano, P. Perry, R. Ashman R. & K. Waite, "The influence of image interactivity upon user engagement when using mobile touch screens," *Comput. Hum. Behav.*, vol. 77, pp. 406-412, 2017.
- [458] Tobii Technology AB (2010). "Tobii Eye Tracking. An introduction to eye tracking and Tobii Eye Trackers," Tobii Technology, 2010. [Online]. Available: <http://acuity-ets.com/downloads/Tobii%20Eye%20Tracking%20Introduction%20Whitepaper.pdf>. [Accessed: June 16, 2017].
- [459] L. T. Kessels and R. A. Ruiter, "Eye movement responses to health messages on cigarette packages," *BMC Public Health*, vol. 12, pp. 352, 2012.
- [460] I. Kersbergen and M. Field, "Alcohol consumers' attention to warning labels and brand information on alcohol packaging: Findings from cross-sectional and experimental studies," *BMC Public Health*, vol. 17, pp. 123, 2017.
- [461] M. Mackert, S.E. Champlin, K.E. Pasch and B.D. Weiss, "Understanding Health Literacy Measurement Through Eye Tracking," *J. Health Commun.*, vol. 18 Suppl. 1, pp. 185-196, 2013.
- [462] A. Poole and L.J. Ball. "Eye tracking in human-computer interaction and usability research: current status and future prospects," in *Encyclopedia of Human-Computer Interaction*, C. Ghaoui, Ed. IGI Global, 2006, pp. 211-219.
- [463] S. B. Bass, T.F. Gordon, R. Gordon and C. Parvanta, "Using eye tracking and gaze pattern analysis to test a "dirty bomb" decision aid in a pilot RCT in urban adults with limited literacy," *BMC Med. Inform. Decis. Mak.*, vol. 16, pp. 67, 2016.
- [464] Diabetes UK, "Understanding diabetes. Your essential guide," Diabetes UK, London, 2011. [Online]. Available: <https://www.diabetes.org.uk/upload/How%20we%20help/catalogue/Understanding%20diabetes%20LG%20PRINT.pdf>. [Accessed: June 16, 2017].

- [465] Tobii Technology AB, "Tobii T/X series Eye Trackers," Tobii Technology AB, 2010. [Online]. Available: <http://www.tobiipro.com/siteassets/tobii-pro/product-descriptions/tobii-pro-tx-product-description.pdf>. [Accessed: June 16, 2017].
- [466] A. Olsen, "The Tobii I-VT Fixation Filter. Algorithm description," Tobii Technology, 2012. [Online]. Available: <http://acuity-ets.com/downloads/Tobii%20I-VT%20Fixation%20Filter.pdf>. [Accessed: June 16, 2017].
- [467] F. C. Bull, C. Holt, M. Kreuter, E. Clark and D. Scharff, "Understanding the effects of printed health education materials: which features lead to which outcomes?" *J. Health Commun.*, vol. 6, pp. 265-280, 3, 2001.
- [468] N. Bol, J.C. van Weert, E.F. Loos, J.C. Romano Bergstrom, S. Bolle and E.M. Smets, "How are online health messages processed? Using eye tracking to predict recall of information in younger and older adults," *J. Health Commun.*, vol. 21, pp. 387-396, 4, 2016.
- [469] N. Bol, J.C. van Weert, H.C. de Haes, E.F. Loos, S. de Heer, D. Sikkels and E.M. Smets, "Using cognitive and affective illustrations to enhance older adults' website satisfaction and recall of online cancer-related information," *Health Commun.*, vol. 29, pp. 678-688, 7, 2014.
- [470] S. Bakhshi, D. A. Shamma and E. Gilbert, "Faces engage us: Photos with faces attract more likes and comments on instagram," in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, 2014, pp. 965-974.
- [471] S. Fletcher-Watson, J.M. Findlay, S.R. Leekam and V. Benson, "Rapid detection of person information in a naturalistic scene," *Perception*, vol. 37, pp. 571-583, 4, 2008.
- [472] E. Birmingham, W.F. Bischof and A. Kingstone, "Saliency does not account for fixations to eyes within social scenes," *Vision Research*, vol. 49, pp. 2992-3000, 24, 2009.
- [473] M. Cerf, J. Harel, W. Einhäuser and C. Koch, "Predicting human gaze using low-level saliency combined with face detection," in *Advances in Neural Information Processing Systems 20: 21st Annual Conference on Neural Information Processing Systems 2007*, J.C. Platt, D. Koller, Y. Singer and S.T. Roweis, Eds. Curran Associates Inc, 2008, pp. 241-248.

- [474] C. Tsai and S. Guan, "Identifying regions of interest in reading an image," *Displays*, vol. 39, pp. 33-41, 2015.
- [475] X. Min, G. Zhai and K. Gu, "Visual attention on human face," in *2015 Visual Communications and Image Processing (VCIP)*, 2015, pp. 1-4.
- [476] C.T. Vu, E.C. Larson and D.M. Chandler, "Visual fixation patterns when judging image quality: Effects of distortion type, amount, and subject experience," in *Image Analysis and Interpretation, 2008. SSIAI 2008. IEEE Southwest Symposium on*, 2008, pp. 73-76.
- [477] E. Birmingham, W.F. Bischof and A. Kingstone, "Why do we look at people's eyes?" *Journal of Eye Movement Research*, vol. 1, 1, 2007.
- [478] N. Diviani, A.L. Dima and P.J. Schulz, "A Psychometric Analysis of the Italian Version of the eHealth Literacy Scale Using Item Response and Classical Test Theory Methods," *J. Med. Internet Res.*, vol. 19, pp. e114, 4, 2017.
- [479] S. Chung and E. Nahm, "Testing Reliability and Validity of the eHealth Literacy Scale (eHEALS) for Older Adults Recruited Online," *Comput. Inform. Nurs.*, vol. 33, pp. 150-156, 4, 2015.

Appendix A

Newest Vital Sign (NVS)

With the NVS assessment tool the participant reads an ice cream container nutritional label and then answers the following questions [359, 362].

NVS Showcard:

Product Description: Ice Cream	
Serving Size:	100ml
Servings per container:	4
NUTRITIONAL INFORMATION	
TYPICAL VALUES	Per 100ml
Energy	1050 kJ
	250 kcal (calories)
Protein	4 g
Carbohydrate	30 g
of which sugars	23 g
Fat	13 g
of which saturates	9 g
of which monounsaturates	0 g
of which polyunsaturates	3 g
of which trans fats	1 g
Fibre	0 g
Sodium	0.05 g
Ingredients: Cream, Skimmed Milk, Sugar, Whole Egg, Stabilisers (Guar Gum), Peanut Oil, Vanilla Extract (0.05%).	

NVS Questions & Answers:

- How many calories (kcal) will you eat if you eat the whole container?
Answer: 1,000 KCAL or 1,000 CALORIES
- If you are advised to eat no more than 60 grams of carbohydrate for dessert, what is the maximum amount of ice cream you could have?
Answer: Two servings (or anything up to 2 servings) OR Half the container (or any amount up to half the container) OR 200 ml (or any amount up to 200 ml).
- Imagine that your doctor advises you to reduce the amount of saturated fat in your diet. You usually have 42 g of saturated fat each day, some of which comes from one serving of ice cream. If you stop eating ice cream, how many grams of saturated fat would you be eating each day?
Answer: 33g
- If you usually eat 2500 calories each day, what percentage of your daily calorie (kcal) intake will you get if you eat one serving of ice cream?
Answer: 1/10 (one tenth) OR 10%

Imagine that you are allergic to the following substances: penicillin, peanuts, latex gloves, and bee stings.

5. Is it safe for you to eat this ice cream?

Answer: no

If 'No' to Q5:

6. Why not?

Answer: Because it contains peanut oil/peanuts/nuts OR Because you might have an allergic reaction

ASK IF answer to Q6 is 'Because you might have an allergic reaction':

7. Why would you have an allergic reaction?

Answer: Because it contains peanut oil/peanuts/nuts

Appendix B

The eHealth Literacy Scale (eHEALS)

The eHEALS comprises questions 3-10. Questions 1 and 2 are supplementary items for use with the eHEALS to understand a participant's interest in using eHealth in general [360].

1. How **useful** do you feel the Internet is in helping you in making decisions about your health?

Not useful at all	Not useful	Unsure	Useful	Very Useful
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. How **important** is it for you to be able to access health resources on the Internet?

Not important at all	Not important	Unsure	Important	Very Important
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. I know **what** health resources are available on the Internet.

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. I know **where** to find helpful health resources on the Internet.

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. I know **how** to find helpful health resources on the Internet.

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. I know **how to use** the Internet to answer my questions about health.

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. I know how to use **the health information** I find on the Internet to help me.

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. I have the skills I need to **evaluate** the health resources I find on the Internet.

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. I can tell **high quality** health resources from **low quality** health resources on the Internet.

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10. I feel **confident** in using information from the Internet to make health decisions.

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Appendix C

Pseudo code for the HCI Parser

Input: A log file which holds multiple records relating to browser interactions during 1 health question

```

for each record in log file
    webPageObj  $\leftarrow$  null
    recordArr[]  $\leftarrow$  tokenise log file record using tab as delimiter
    dateStr  $\leftarrow$  recordArr[n]
    tokenise dateStr  $\rightarrow$  day, month, year then convert to integer values
    timeStr  $\leftarrow$  recordArr[n]
    tokenise timeStr  $\rightarrow$  second, minute, hour then convert to integer values
    webPageObj.timestamp  $\leftarrow$  use integer values to construct Java Calendar object
    webPageObj.eventType  $\leftarrow$  recordArr[n]
    webPageObj.taskType  $\leftarrow$  recordArr[n]
    URLStr  $\leftarrow$  recordArr[n]
    if URLStr is valid URL then
        webPageObj.URL  $\leftarrow$  URLStr
        webPageObj.domain  $\leftarrow$  extract domain from webPageObj.URL
        if webPageObj.URL has query strings then
            for each query string
                webPageObj.queryStr[n]  $\leftarrow$  extract query string
                    from webPageObj.URL
            end for
        end if
    end if

    if webPageObj.eventType is “submittedAnswerText” then
        webPageObj.answer = recordArr[n]
    else
        if webPageObj.taskType is “posttask”
            && webPageObj.eventType is question
            no. then
                webPageObj.posttaskanswer = recordArr[n]
            end if
        end if

    webPagesArr[n]  $\leftarrow$  webPageObj
    webPagesArr[n-1].duration  $\leftarrow$  webPagesArr[n].timestamp –
        webPagesArr[n-1].timestamp

end for

```

Appendix D

Ontology Engineering Survey

The ontology will capture information about a patient such as personal information and health and fitness characteristics. This information will be used to tailor the education to the patient's needs.

Cardinality refers to the number of pieces of data that are related. For example if a patient can have only 1 Body Mass Index (BMI) captured in the ontology then we would say that the cardinality of the relationship from patient to BMI is *exactly 1*. However if a patient must be treated by a doctor, and can be treated by more than 1 doctor, then the cardinality of the relationship from patient to doctor is *at least 1*.

1. Can you suggest any health characteristics that should be captured for a patient?
How many pieces of information would be captured for each patient?

The ontology will be used to provide information about ways to incorporate more activity in a patient's lifestyle. This can include sports or everyday activities. A class called **Activity** could be modelled to capture activities. It could contain the subclasses **Light intensity**, **Moderate intensity** and **Vigorous intensity** to further categorise activities. This class hierarchy is illustrated in Figure 1.



Light intensity activities

- No noticeable change in breathing patterns.

Moderate intensity activities

- You will breathe harder and your heart will beat faster.
- You should be able to carry on a conversation but cannot sing a song.

Vigorous intensity activities

- You will breathe much harder and your heart will beat rapidly.
- You may find it difficult to carry on a conversation.

2. For each activity indicate the class to which it should belong (Aerobics, Basketball, Canoeing, Climbing Stairs, Cycling Level Ground, Cycling Fast, Cycling Over

Hills, Doubles Tennis, Fishing, Golf, Gymnastics, Ice Hockey, Jogging, Lifting Weights, Martial Arts, Rollerblading, Rugby, Running, Singles Tennis, Skateboarding, Skipping, Sprinting, Strolling, Swimming, Walking Fast, Water Aerobics, Volleyball, Yoga).

The ontology will capture the different physical activity levels that can be achieved by an adult. These activity levels could be captured in a class called **Physical Activity Level**. Class Physical Activity Level could contain subclasses to further categorise activity levels. For example, sedentary behaviour involves activities that do not increase energy expenditure much above resting levels, therefore **Sedentary** could be a subclass of Physical Activity Level.

3. Can you suggest any subclasses of class Physical Activity Level?

The ontology will capture information about the health conditions Diabetes and Obesity.

- **Obesity**: Condition in which excess fat has accumulated in the body.
- **Diabetes**: Condition characterised by lack of insulin needed for sugar metabolism.
- **Type 1 Diabetes** develops when the body is unable to produce any insulin.
- **Type 2 Diabetes** develops when the body can make insulin, but not enough.

The ontology could contain 5 classes to represent these conditions; **Health Condition**, **Obesity**, **Diabetes**, **Type 1 Diabetes** and **Type 2 Diabetes**. These could be organised in the following class hierarchy,

Figure 2



4. Can you suggest different classes (and subclasses) to represent the health conditions Obesity and Diabetes?

The ontology will capture information about the symptoms of Obesity and Diabetes. These could be organised in a class hierarchy similar to Figure 3.



5. Can you suggest different names for these classes (or a different way to organise symptoms)?
6. Can you suggest any symptoms and indicate how these should be classified in the ontology?
7. Can you suggest any other information that should be captured about any of the symptoms of Diabetes or Obesity?

The ontology will capture information about treatments for Diabetes and Obesity. Treatments could be organised in a class hierarchy similar to Figure 4.



8. Can you suggest different names for these classes (or a different way to organise treatments)?
9. Can you suggest any non-medication treatments (e.g. essential health checks) and indicate how these should be classified in the ontology?

The ontology will capture information about the Diabetes treatment Insulin. Class **Diabetes Treatments** could have a subclass called **Insulin Types** to capture information about the different types of Insulin.

10. Can you suggest a different name for class Insulin Types?
11. Can you suggest any subclasses of class Insulin Types?
12. Can you suggest any other information that should be captured about Insulin?

The ontology will also capture information about non-Insulin medications. Class **Diabetes Treatments** could have a subclass called **Diabetes Medication** to capture information about non-Insulin medications.

13. Can you suggest a different name for this class?
14. Can you suggest any subclasses of class Diabetes Medication?
15. Can you suggest any other information that should be captured about non-Insulin medication?

The ontology will capture information about the complications associated with Obesity and Diabetes. These could be organised in a class hierarchy similar to Figure 5.



16. Can you suggest different names for these classes (or a different way to organise complications)?
17. Can you suggest any complications and indicate how these should be classified in the ontology?
18. Can you suggest any other information that should be captured about any of the complications you listed?

Appendix E

User Engagement Scale (UES)

The User Engagement Scale is a questionnaire that focuses on dimensions of a user's experience during interaction with a system [455, 456].

Aesthetics (AE)

- The patient education is aesthetically appealing.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- The patient education is attractive.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- I liked the graphics and images used in the patient education.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- The patient education appealed to my visual senses.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- The layout of the patient education was visually pleasing.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Perceived Usability (PUs)

- I found the patient education confusing to use.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- I felt frustrated while searching for information in the patient education.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- I felt annoyed while searching for information in the patient education.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- I felt discouraged while searching for information in the patient education.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- Searching for information in the patient education was mentally taxing.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- Searching for information in the patient education was demanding.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>